R. Courant · H. Robbins



CE ESTE MATEMATICA?

Traducere de S. Teleman

Coperta și supracoperta: D. Petrescu

R. Courant • H. Robbins

CE ESTE MATEMATICA?

Expunere elementară a ideilor și metodelor



Editura Științifică

BUCUREȘTI, 1969

PREFAȚĂ LA PRIMA EDIȚIE

Timp de peste două mii de ani, o oarecare familiaritate cu matematica a fost privită ca parte indispensabilă a înzestrării intelectuale a oricărei persoane culte. Astăzi, locul tradițional al matematicii în educație se află într-un grav pericol. Din nefericire, vina o poartă și reprezentanții profesioniști ai matematicii. Predarea matematicii a degenerat uneori într-o dexteritate seacă de rezolvare a problemelor, care poate dezvolta îndemînarea formală, dar nu duce la o înțelegere reală sau la o mai mare independență intelectuală. Cercetarea matematică a căpătat o tendință de supraspecializare și supraaccentuare a abstractizării. Aplicațiile și legăturile cu alte domenii au fost neglijate. Aceste condiții nu justifică însă cîtuși de puțin o atitudine defensivă. Dimpotrivă, trebuie să apară, și chiar apare, o reacție opusă din partea acelora care sînt conștienți de valoarea disciplinei intelectuale. Profesorii, studenții și publicul educat cer o reformă constructivă și nu resemnarea pe drumul rezistenței minime. Scopul este o reală înțelegere a matematicii, ca un tot organic și ca bază a gîndirii și acțiunii științifice.

Cîteva cărți excelente cu caracter biografic și istoric și unele scrieri atractive de popularizare au stimulat interesul general aflat în stare latentă. Însă cunoașterea nu poate fi obținută numai prin metode indirecte. Înțelegerea matematicii nu poate fi transmisă prin distracție lipsită de efort, tot așa cum nici educația muzicală nu poate fi realizată prin citirea celor mai strălucite articole despre muzică de către aceia care nu au ascultat niciodată muzica cu intensitate. Este necesar contactul real cu conținutul matematicii vii. Totuși, aspectele tehnice și ocolurile trebuie evitate, iar prezentarea matematicii ar trebui să fie tot atît de lipsită de rutină, ca și de dogmatismul care interzice dezvăluirea motivelor sau scopurilor și care este un obstacol incorect pus în fața efortului cinstit. Este posibil să înaintăm pe un drum drept, de la primele elemente pînă la rezultatele esențiale, de unde pot fi privite substanța și forțele care dau impuls matematicii moderne.

Această carte constituie o încercare în această direcție. În măsura în care ea presupune doar cunoașterea matematicii de liceu, ea poate fi privită ca o lucrare de popularizare. Însă ea nu face concesii tendinței periculoase de eliminare a oricărui efort. Ea pretinde un anumit grad de maturitate intelectuală și dorința de a gîndi cu propria minte. Cartea este scrisă pentru începători și învățați, pentru studenți și profesori, pentru filozofi și ingineri, pentru elevi și publicul larg. Este posibil ca aceasta să fie o intenție prea ambițioasă. Sub presiunea altor obligații a trebuit să facem un oarecare compromis la publicarea cărții, care urmează după mulți ani de pregătire, dar înainte ca ea să fie într-adevăr terminată. Critica și sugestiile vor fi de aceea binevenite.

În timp ce răspunderea pentru planul și filozofia acestei lucrări revine subsemnatului, orice merite pe care ea ar putea să le aibă trebuie să fie împărțite cu Herbert Robbins. Îndată ce el s-a asociat la redactarea ei, a făcut din aceasta propria sa cauză, și colaborarea lui a jucat un rol decisiv în aducerea lucrării la forma actuală.

Datorăm mulțumiri recunoscătoare multor prieteni pentru ajutorul dat. Discuțiile cu Niels Bohr, Kurt Friedrichs și Otto Neugebauer au influențat poziția mea asupra problemelor cu caracter filozofic și istoric; Edna Kramer a contribuit cu o critică constructivă, făcută din punctul de vedere al pedagogului; David Gilbarg a pregătit primele note de curs, în care se află originea cărții; Ernest Courant, Norman Davids, Charles de Prima, Alfred Horn, Herbert Mintzer, Wolfgang Wazow și alții au participat la lucrările nesfîrșite de scriere și rescriere a manuscrisului și au contribuit considerabil la îmbunătățirea amănuntelor; Donald Flanders a făcut multe sugestii prețioase și a verificat manuscrisul pentru tipar; John Knudsen, Hertha von Gumppenberg, Irving Ritter și Otto Neugebauer au pregătit desenele; H. Whitney a contribuit la culegerea exercițiilor din apendice. Comisia generală de educație a Fundației Rockefeller a sprijinit cu generozitate dezvoltarea cursurilor și notelor care au devenit apoi baza acestei cărți. Mai datorăm mulțumiri editurii Waverly Press, în special d-lui Grover C. Orth pentru munca lor extrem de competentă, și editurii Oxford University Press, în special d-lui Philip Vaudrin și d-lui W. Oman pentru inițiativa lor încurajatoare și pentru colaborare.

New Rochelle, New York, 22 august 1941

R. COURANT

PREFAȚĂ LA EDIȚIILE A DOUA, A TREIA ȘI A PATRA

În cursul ultimilor ani, forța evenimentelor a dus la o cerere crescută de instruire și informare matematică. Acum, mai mult decît oricînd, există pericolul frustrării și deziluziei dacă studenții și profesorii nu încearcă să privească dincolo de formalismul și tehnica matematică și nu încearcă să sesizeze adevărata esență a matematicii. Această carte a fost scrisă tocmai pentru acești studenți și profesori, și primirea de care s-a bucurat prima ediție încurajează autorii, în speranța că ea va fi folositoare.

Critica primită din partea a numeroși cititori a dus la multe îndreptări și îmbunătățiri. Pentru ajutorul generos dat la pregătirea celei de-a patra ediții, datorăm mulțumiri cordiale doamnei Natascha Artin.

R. COURANT

New Rochelle, N. Y., 18 martie 1943
10 octombrie 1945 — 28 octombrie 1947

CUM TREBUIE FOLOSITĂ CARTEA

Cartea este scrisă într-o ordine sistematică, dar aceasta nu înseamnă cîtuşi de puţin că cititorul este obligat să o parcurgă pagină după pagină, capitol după capitol. Capitolele sînt destul de independente unul faţă de altul. Adesea, începutul paragrafului este uşor accesibil, dar apoi, drumul urcă treptat, devenind mai abrupt spre sfîrşitul capitolului şi în supliment. De aceea cititorul care doreşte mai degrabă o informare generală decît o cunoaştere amănunţită se poate mulţumi cu alegerea materialului, evitînd discuţiile mai amănunţite.

Studentul care dispune de cunoştinţe matematice reduse va avea de făcut şi el o alegere. Asteriscul, sau caracterul petit, indică părţi care pot fi omise la prima citire, fără a dăuna în mod serios înţelegerii părţilor care urmează. Mai mult, nu se va pierde nimic dacă studiul cărţii se va restrînge la acele paragrafe sau capitole care prezintă cel mai mare interes pentru cititor. Cea mai mare parte a exerciţiilor nu au ca scop formarea rutinei; cele mai grele sînt însemnate cu asterisc. Cititorul nu ar trebui să se alarmeze dacă nu poate rezolva multe din acestea.

Profesorii de liceu pot găsi subiecte utile pentru cercurile de elevi, în capitolele referitoare la construcţiile geometrice şi la maxime şi minime.

Sperăm că această carte va fi utilă studenţilor, începînd cu cei mai tineri pînă la cei mai avansaţi, ca şi persoanelor de diverse profesiuni care manifestă un interes real pentru ştiinţă. Mai mult, ea poate servi ca bază pentru lecţii facultative asupra conceptelor fundamentale ale matematicii. Capitolele III, IV şi V ar putea fi folosite pentru un curs de geometrie, în timp ce capitolele VI şi VIII formează, la un loc, o prezentare de sine stătătoare a

analizei matematice, în care accentul este pus mai degrabă pe înţelegere, decît pe rutină. Ele ar putea folosi ca text introductiv profesorilor care vor să aducă contribuţii active la completarea noţiunilor, în concordanţă cu nevoile specifice şi mai ales să le îmbogăţească cu diferite exemple. Numeroase exerciţii răspîndite în întreaga lucrare şi o culegere suplimentară de la sfîrşitul ei vor uşura folosirea cărţii în clasă.

Sperăm chiar că şi specialistul va fi interesat în anumite amănunte şi discuţii elementare, care conţin germenele unei dezvoltări mai largi.

CE ESTE MATEMATICA?

Matematica reflectă voința activă, rațiunea contemplativă și dorința de perfecțiune estetică. Elementele ei de bază sînt logica și intuiția, analiza și construcția, generalul și concretul. Cu toate că diferitele tradiții pot sublinia aspecte diferite, doar interacțiunea acestor forțe polare și lupta de a le sintetiza constituie viața, utilitatea și valoarea supremă a științei matematice.

Fără îndoială, orice dezvoltare matematică își are rădăcinile în necesități mai mult sau mai puțin practice. Dar o dată apărută sub presiunea aplicațiilor necesare, ea capătă în mod inevitabil un impuls din ea însăși și depășește limitele utilității imediate. Această tendință de trecere de la știința aplicată la cea teoretică se observă atît în istoria antică, cît și în zilele noastre: este suficient să avem în vedere contribuția adusă matematicii moderne de ingineri și fizicieni.

Primele documente matematice provin din Orient, unde cu aproximativ 2000 de ani î.e.n. babilonenii au strîns un imens material, pe care astăzi l-am considera ca aparținînd algebrei elementare. Și totuși, ca știință în sensul modern al cuvîntului, matematica apare mai tîrziu, pe pămîntul Greciei, în cel de-al cincilea și al patrulea secol î.e.n. Contactul din ce în ce mai strîns dintre Orient și greci, care a început pe timpul Imperiului persan și a atins un apogeu în perioada care a urmat expedițiilor lui Alexandru cel Mare, a făcut cunoscute grecilor realizările matematicii și astronomiei babilonene. Matematica a fost supusă curînd discuției filozofice care înflorea în cetățile grecești. Astfel, gînditorii greci au devenit conștienți de marile dificultăți inerente noțiunilor matematice de continuitate, mișcare și infinit, ca și de cele legate de problema măsurării unor cantități arbitrare prin unități date. Printr-un efort admirabil ei au făcut față acestei provocări și rezultatul, teoria lui Eudoxus a continuului geometric, este o realizare care a fost egalată doar după mai bine de 2000 de ani, de teoria modernă a numerelor iraționale. Tendința deductiv-axiomatică din matematică își are originea în vremea lui Eudoxus și a fost cristalizată în *Elementele* lui Euclid.

Însă, în timp ce tendința teoretic-axiomatică a matematicii grecești rămîne una din caracteristicile ei importante și a exercitat o influență enormă, nu exagerăm subliniind faptul că aplicația și legătura cu realitatea fizică au jucat un rol tot atît de important în matematica antichității și că foarte adesea era preferat un stil de prezentare mai puțin rigid decît acela al lui Euclid.

Este posibil ca primele descoperiri ale dificultăților legate de cantitățile „incomesurabile" să fi abătut pe greci de la dezvoltarea artei calculului numeric, realizată anterior în Orient. În schimb, ei au deschis un drum în desișul geometriei axiomatice pure. Astfel a început unul din ciudatele ocoluri din istoria științei și este posibil să se fi pierdut o ocazie importantă. Timp de aproape două mii de ani, prestigiul tradiției geometrice grecești a întîrziat evoluția inevitabilă a noțiunii de număr și a calculului algebric, care mai tîrziu au fost puse la baza științei moderne.

După o perioadă de pregătire lentă, revoluția din matematică și știință și-a început faza viguroasă în secolul al XVII-lea, cu geometria analitică și cu calculul diferențial și integral. În timp ce geometria greacă și-a menținut un loc important, idealul grec al cristalizării axiomatice și al deducției sistematice a dispărut în secolele XVII și XVIII. Raționamentul logic precis, care pornește de la definiții clare și axiome necontradictorii „evidente", părea lipsit de importanță noilor pionieri ai științei matematice. Într-o adevărată orgie a ghicirii intuitive, a raționamentului convingător întrețesut cu misticism lipsit de sens, cu o încredere oarbă în puterea supraomenească a procedeului formal, ei au cucerit o lume matematică de o imensă bogăție. Treptat, extazul progresului a făcut loc unui spirit de autocontrol critic. În secolul al XIX-lea, necesitatea inevitabilă de consolidare a științei și dorința unei mai mari siguranțe la extinderea învățămîntului superior, care a fost promovată de revoluția franceză, au obligat la o revizuire a fundamentelor noii matematici; în particular, a calculului diferențial și integral și a noțiunii fundamentale de limită. Astfel, secolul al XIX-lea a devenit nu numai o perioadă de noi succese, dar a fost caracterizat și printr-o revenire fericită la idealul clasic al preciziei și demonstrației riguroase. În această privință, el a depășit chiar modelul științei grecești. Încă o dată pendulul a oscilat spre puritatea logică și abstracție. În momentul de față, se pare că ne mai aflăm în această perioadă, cu toate că este de dorit ca separarea nefericită dintre matematica pură și aplicațiile ei vitale, inevitabilă, după cît se pare, într-o perioadă de revedere critică, să fie urmată de o eră de unitate mai strînsă. Forța internă recîștigată și, mai ales, enorma simplificare obținută printr-o înțelegere mai clară au făcut posibilă astăzi stăpînirea teoriei matematice, fără a pierde din vedere aplicațiile ei. Restabilirea unei legături organice între știința pură și cea aplicată și a unui echilibru sănătos între generalitatea abstractă și particularitatea concretă plină de culoare ar putea fi datoria majoră a matematicii în viitorul imediat.

Nu este acesta locul pentru a face o analiză filozofică sau psihologică detaliată a matematicii. Ar trebui să subliniem numai cîteva momente. Pare să existe un mare pericol în supraaccentuarea caracterului deductiv-axiomatic al matematicii, dominant astăzi. Este adevărat, elementul invenției constructive, al intuiției directoare și motivatoare, poate scăpa unei formulări filozofice simple; însă el rămîne sursa oricărei realizări matematice, chiar în cele mai abstracte domenii. Dacă scopul este forma deductivă cristalizată, intuiția și construcția sînt cel puțin forțele conducătoare. În afirmația că matematica nu este decît un sistem de concluzii trase din definiții și axiome, care trebuie să fie consistente, iar restul este produsul liberei fantezii a matematicianului, se ascunde o amenințare serioasă la însăși viața științei. Dacă această descriere ar fi corectă, matematica nu ar putea atrage nici o persoană inteligentă. Ea ar fi un simplu joc cu definiții, reguli și silogisme, fără nici un principiu sau scop. Ideea că intelectul uman poate crea sisteme axiomatice semnificative, după propriile sale capricii, este un semiadevăr amăgitor. Doar subordonată responsabilității față de întregul organic, doar condusă de necesitatea intrinsecă, mintea liberă poate să obțină rezultate valoroase din punct de vedere științific.

Deşi tendința contemplativă a analizei logice nu reprezintă întreaga matematică, ea a dus la o înțelegere mai profundă a faptelor matematice și a interdependenței lor și la o înțelegere mai clară a esenței noțiunilor matematice. Din ea s-a dezvoltat punctul de vedere modern asupra matematicii ca model universal al metodei științifice aplicative.

Oricare ar fi punctul nostru de vedere filozofic, pentru toate scopurile cercetării științifice, un obiect se dezvăluie prin totalitatea relațiilor posibile față de subiectul sau instrumentul care îl observă. Desigur, simpla percepție nu constituie o cunoaștere și o pătrundere a naturii lui; ea trebuie să fie coordonată și interpretată în legătură cu o entitate fundamentală, „un lucru în sine", care nu este obiectul observației fizice directe, ci aparține metafizicii. Totuşi, pentru procedeul științific, este important să se îndepărteze elementele cu caracter metafizic și să considerăm întotdeauna faptele observabile ca sursă ultimă a noțiunilor și construcțiilor. Renunțarea la scopul înțelegerii „lucrului în sine", al cunoașterii „adevărului ultim", al dezvăluirii esenței celei mai profunde a lumii, poate fi o suferință psihologică pentru entuziaștii naivi, însă, ea reprezintă de fapt una dintre cele mai fructuoase cotituri din gîndirea modernă.

Unele dintre cele mai mari realizări din fizică s-au obținut ca răsplată pentru aderarea curajoasă la principiul eliminării metafizicii. Cînd Einstein a încercat să reducă noțiunea de „evenimente simultane care se petrec în locuri diferite" la fenomene observabile, cînd el a demascat ca fiind prejudecată metafizică credința că această noțiune trebuie să aibă o semnificație științifică în sine, el a găsit cheia teoriei sale, a relativității. Cînd Niels Bohr și elevii săi au analizat faptul că orice observație fizică trebuie să fie înso-

țită de un efect al instrumentului observator asupra obiectului observat, a devenit clar că fixarea simultană precisă a poziției și vitezei unei particule nu este posibilă din punctul de vedere al fizicii. Consecințele vaste ale acestei descoperiri, încadrate în teoria modernă a mecanicii cuantice, sînt familiare acum fiecărui fizician. În secolul al XIX-lea predomina ideea că forțele mecanice și mișcările particulelor în spațiu sînt lucruri în sine, în timp ce electricitatea, lumina și magnetismul ar trebui să fie reduse la fenomene mecanice sau ,,explicate'' ca fenomene mecanice, tot așa cum s-a făcut cu căldura. A fost inventat ,,eterul'', ca mediu ipotetic, capabil de mișcări mecanice neexplicate pe deplin care ne apar ca lumină sau electricitate. Încetul cu încetul, s-a recunoscut că eterul este în mod necesar inobservabil, că această noțiune aparține metafizicii și nu fizicii. Cu supărare în anumite cercuri, cu ușurare în altele, explicația mecanică dată luminii și electricității, și cu ele și eterul, au fost abandonate în cele din urmă.

O situație similară, chiar mai accentuată, există și în matematică. Din cele mai vechi timpuri matematicienii au considerat obiectele lor, ca de pildă numere, puncte etc., ca substanțe oarecare, lucruri în sine. Deoarece aceste entități refuzau întotdeauna să se supună încercărilor de a le descrie în mod adecvat, încetul cu încetul, matematicienilor din secolul al XIX-lea le-a apărut cu tot mai multă claritate faptul că problema semnificației acestor obiecte ca lucruri substanțiale este lipsită de sens în matematică. Afirmațiile matematice în care intră acești termeni nu se referă nicidecum la realitatea fizică; ele stabilesc doar relații existente între ,,obiecte nedefinite'' matematic și reguli care guvernează operațiile cu ele. Ce sînt ,,într-adevăr'' punctele, dreptele, numerele, este un lucru care nu poate și nu trebuie să fie discutat în matematică. Ceea ce are importanță, și ceea ce corespunde faptului ,,verificabil'', este structura și relația, faptul că două puncte determină o dreaptă, că numerele se combină în conformitate cu anumite reguli, pentru a forma alte numere etc. O intuire clară a necesității dematerializării noțiunilor matematice elementare a fost unul dintre cele mai importante și fructuoase rezultate ale dezvoltării axiomatice moderne.

Din fericire, mințile creatoare uită credințele filozofice dogmatice, ori de cîte ori aderarea la ele ar fi o piedică în calea descoperirilor constructive. Pentru specialiști, ca și pentru amatori, nu filozofia, ci experiența activă în însăși matematică este singura care poate da răspuns întrebării: ce este matematica?

Capitolul I

NUMERELE NATURALE

INTRODUCERE

Numărul este fundamentul matematicilor moderne. Dar ce este numărul? Ce înseamnă că $\frac{1}{2} + \frac{1}{2} = 1$, $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ sau că $(-1)(-1) = 1$? În școală învățăm tehnica operării cu fracții și numere negative, dar pentru o reală înțelegere a modului cum a fost construit sistemul de numere nu este de ajuns să ne limităm la cunoștințe elementare, ci trebuie să mergem mai departe. În antichitate, grecii au pus la baza matematicii create de ei conceptele geometrice de punct și dreaptă; principiul dominant al matematicii moderne constă în faptul că toate propozițiile matematice trebuie să se reducă, în cele din urmă, la propoziții referitoare la *numerele naturale*, 1, 2, 3, ... „Dumnezeu a creat numerele naturale; restul este opera omului". Cu aceste cuvinte, Leopold Kronecker (1823—1891) a indicat ... „fundamentul solid" pe care se poate construi structura matematicii.

Create de intelectul uman pentru a număra obiectele din diferite colecții, numerele nu se referă la caracteristicile individuale ale obiectelor numărate. Astfel, numărul 6 este o abstracție a tuturor colecțiilor reale formate din 6 obiecte; el nu depinde de nici o proprietate specifică a acestor obiecte sau de simbolurile folosite. Caracterul abstract al ideii de număr devine clar numai la un stadiu relativ avansat al dezvoltării intelectuale. Pentru copii, numerele rămîn întotdeauna legate de obiecte palpabile, ca de pildă degete sau pietricele; în limbile diferitelor triburi, numerele au încă un sens numeric concret, folosindu-se cuvinte diferite pentru numărarea obiectelor de diferite tipuri.

Din fericire, matematicianul nu trebuie să se ocupe de problema filozofică a trecerii de la colecții de obiecte concrete la conceptul abstract de număr. De aceea, vom accepta numerele naturale ca fiind date împreună cu cele două operații fundamentale care le determină: adunarea și înmulțirea.

1. Legile aritmeticii

Teoria matematică a *numerelor naturale*, sau a *întregilor pozitivi*, este cunoscută sub numele de *aritmetică*. Ea se bazează pe faptul că adunarea și înmulțirea întregilor sînt guvernate de cîteva legi. Pentru a enunța aceste legi în toată generalitatea lor, nu putem utiliza simboluri de genul 1, 2, 3 care se referă la anumiți întregi. Propoziția

$$1 + 2 = 2 + 1$$

este doar un exemplu particular al legii generale, în virtutea căreia suma a doi întregi este aceeași, indiferent de ordinea în care considerăm aceste numere. Prin urmare, cînd dorim să exprimăm faptul că o anumită relație între întregi este valabilă, indiferent de valorile întregilor particulari care intervin, vom nota întregii simbolic, prin literele a, b, c, ... Cu această convenție putem enunța cinci legi fundamentale ale aritmeticii, cu care cititorul este familiarizat:

1) $a + b = b + a$, 2) $ab = ba$,

3) $a + (b + c) = (a + b) + c$, 4) $a(bc) = (ab)c$,

5) $a(b + c) = ab + ac$.

Primele două, legile *comutativității* adunării și înmulțirii, afirmă că se poate schimba ordinea elementelor care intervin în adunare sau înmulțire. A treia, legea *asociativității* adunării, afirmă că prin adunarea a trei numere se obține același rezultat, fie că adunăm la primul suma celui de-al doilea și al treilea, fie că adunăm celui de-al treilea suma primului și a celui de-al doilea. A patra, este legea *asociativității* înmulțirii. Ultima, legea *distributivității*, exprimă faptul că pentru a înmulți o sumă cu un întreg putem înmulți fiecare termen al sumei cu acest întreg și apoi să adunăm produsele.

Aceste legi ale aritmeticii sînt foarte simple și pot părea evidente. Dar trebuie să remarcăm că ele pot să nu fie aplicabile altor entități. De exemplu, dacă a și b nu reprezintă numere, ci substanțe chimice, și dacă „adunarea" este folosită în sensul obișnuit al cuvîntului, este evident că legea comutativității nu este valabilă întotdeauna. Într-adevăr, dacă, de exemplu, adăugăm acid sulfuric peste o cantitate de apă obținem o soluție diluată, în timp ce adăugarea unei picături de apă peste acid sulfuric pur poate provoca un accident experimentatorului. Cu ajutorul unor exemple de același gen se poate arăta că în acest tip de „aritmetică" chimică, legile asociativității și distributivității pot să nu mai fie valabile. Astfel, ne putem imagina

tipuri dearitm etică în care una sau mai multe din legile 1)—5) nu sînt valabile. A stfel de sisteme au fost într-adevăr studiate de matematica modernă.

Un model concret pentru conceptul abstract de număr întreg va indica baza pe care se sprijină legile 1)—5). În loc de a folosi simbolurile numerice obișnuite 1, 2, 3 etc., să notăm întregul care dă numărul de obiecte dintr-o



Fig. 1. Adunarea

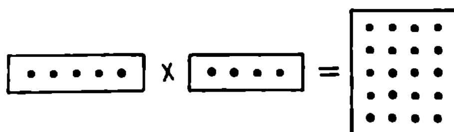


Fig. 2. Înmulțirea

anumită colecție (de exemplu, colecția merelor aflate într-un anumit pom) printr-o mulțime de puncte așezate într-o cutie dreptunghiulară, astfel încît fiecărui obiect să-i corespundă un punct. Operînd cu aceste cutii putem cerceta legile aritmeticii numerelor întregi. Pentru a aduna doi întregi a și b, așezăm cutiile corespunzătoare cap la cap și îndepărtăm despărțitura.

Pentru a înmulți pe a cu b, așezăm punctele celor două cutii în rînduri și formăm o nouă cutie cu a linii și b coloane de puncte. Acum este clar că regulile 1)—5) exprimă proprietăți intuitiv evidente ale acestor operații cu cutii.

Pe baza definiției adunării a doi întregi putem defini relația de *inegalitate.* Fiecare din propozițiile echivalente, $a < b$ (se citește „a este mai mic decît b”) și $b > a$ (se citește „b este mai mare decît a”) înseamnă că cutia b

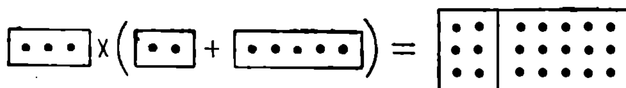


Fig. 3. Legea distributivității

poate fi obținută din cutia a prin adăugarea unei a treia cutii c, aleasă astfel încît $b = a + c$. Dacă se întîmplă acest lucru scriem

$$c = b - a,$$

ceea ce definește operația de *scădere.*

Adunarea și scăderea se numesc *operații inverse*, deoarece dacă adunarea întregului d la întregul a este urmată de scăderea întregului d, rezultatul este întregul inițial a:

$$(a + d) - d = a.$$



Eig. 4. Scăderea

Trebuie să remarcăm că întregul $b - a$ a fost definit numai dacă $b > a$. Interpretarea simbolului $b - a$ ca *întreg negativ*, în cazul în care $b < a$ va fi discutată mai tîrziu (p. 70 și urm.).

Este convenabil, de obicei, să folosim una din notațiile $b \geqslant a$ (se citește „b este mai mare sau egal cu a") sau $a \leqslant b$ (se citește „a este mai mic sau egal b") pentru a exprima negația propoziției $a > b$. Astfel, $2 \geqslant 2$ și $3 \geqslant 2$.

Putem să extindem încă puțin domeniul întregilor pozitivi pe care i-am reprezentat prin cutii cu puncte, întroducînd întregul *zero*, reprezentat printr-o cutie complet goală. Dacă notăm cutia goală prin simbolul obișnuit 0, atunci conform definiției date adunării și înmulțirii avem

$$a + 0 = a,$$

$$a \cdot 0 = 0,$$

pentru orice întreg a. Într-adevăr, $a + 0$ înseamnă adăugarea unei cutii goale cutiei a, în timp ce, $a \cdot 0$ înseamnă o cutie fără coloane; adică, o cutie goală. De aceea, este natural să extindem definiția scăderii punînd

$$a - a = 0,$$

pentru orice întreg a. Acestea sînt proprietățile aritmetice caracteristice ale lui zero.

Modele geometrice, de genul acestor cutii cu puncte sau ca vechiul abac, au fost utilizate pe larg pentru calcule numerice pînă tîrziu în evul mediu, și numai încetul cu încetul ele au fost înlocuite prin metode simbolice cu mult superioare, bazate pe sistemul zecimal.

2. Reprezentarea întregilor

Trebuie să facem cu grijă distincție între un întreg și simbolul prin care este reprezentat (de exemplu 5, V, ... etc.) În sistemul nostru zecimal, zero și primele nouă numere naturale întregi sînt reprezentate prin cifrele

0, 1, 2, 3, ..., 9. Un întreg mai mare ca, de pildă, „trei sute şaptezeci şi doi" se exprimă sub forma

$$300 + 70 + 2 = 3 \cdot 10^2 + 7 \cdot 10 + 2,$$

şi este notat în sistemul zecimal prin simbolul 372. Important este aici faptul că semnificaţia fiecăreia din cifrele 3, 7, 2 depinde de ***poziţia*** lor pe locul unităţilor, zecilor sau sutelor. Folosind această notaţie poziţională putem reprezenta orice număr natural, utilizînd numai cele zece cifre în diferite combinaţii. Regula generală este de a exprima un întreg sub forma

$$z = a \cdot 10^3 + b \cdot 10^2 + c \cdot 10 + d,$$

unde simbolurile a, b, c, d sînt întregi de la zero la nouă. Numărul z este reprezentat în acest caz prin simbolul prescurtat $abcd$. Remarcăm în treacăt că coeficienţii d, c, b, a sînt resturile obţinute prin împărţiri succesive ale lui z prin 10. Astfel,

	Cît	Rest
372 : 10	37	2
37 : 10	3	7
3 : 10	0	3

Cu ajutorul expresiei date mai sus numărului z se pot reprezenta numai întregi mai mici decît zece mii, deoarece întregii mai mari ar necesita cinci sau un număr mai mare de cifre. Dacă z este un întreg, cuprins între zece mii şi o sută de mii, îl putem exprima sub forma

$$z = a \cdot 10^4 + b \cdot 10^3 + c \cdot 10^2 + d \cdot 10 + e,$$

şi îl putem reprezenta simbolic prin $abcde$. O propoziţie similară este valabilă pentru întregi cuprinşi între o sută de mii şi un milion etc. Este foarte important să dispunem de o metodă care să ne permită să exprimăm rezultatul la care ajungem în toată generalitatea sa cu ajutorul unei singure formule. Putem atinge acest scop dacă notăm diferiţii coeficienţi e, d, c, ... printr-o singură literă a cu diferiţi indici, a_0, a_1, a_2, a_3, ..., şi dacă indicăm faptul că puterile lui zece pot fi oricît de mari, reprezentînd cea mai mare putere a lui 10 nu prin 10^3 sau 10^4, ca în exemplele de mai sus, ci prin 10^n, unde n este un întreg arbitrar. Atunci, orice număr întreg z poate fi pus în sistemul zecimal sub forma

$$z = a_n \cdot 10^n + a_{n-1} \cdot 10^{n-1} + \ldots + a_1 \cdot 10 + a_0, \tag{1}$$

şi scris prescurtat prin simbolul

$$a_n a_{n-1} a_{n-2} \ldots a_1 a_0.$$

Ca şi în cazul particular de mai sus, observăm că $a_0, a_1, a_2, \ldots, a_n$ sînt tocmai resturile succesive obţinute prin împărţirea repetată a lui z prin 10.

În sistemul zecimal numărul zece joacă un rol deosebit, servind drept bază a sistemului. Profanul poate să nu remarce că alegerea lui zece nu este esenţială şi că orice întreg mai mare decît unu poate fi ales drept bază. De exemplu, ar putea fi folosit un sistem *septimal* (cu baza 7). Într-un astfel de sistem, un întreg ar fi exprimat sub forma

$$(2) \qquad b_n \cdot 7^n + b_{n-1} \cdot 7^{n-1} + \ldots + b_1 \cdot 7 + b_0,$$

unde coeficienţii b sînt cifre de la zero la şase, şi ar fi notat prescurtat prin simbolul

$$b_n b_{n-1} \ldots b_1 b_0.$$

Astfel, ,,o sută nouă'' ar fi reprezentat în sistemul septimal prin simbolul 214, care înseamnă

$$2 \cdot 7^2 + 1 \cdot 7 + 4.$$

Ca exerciţiu, cititorul poate verifica faptul că regula generală pentru a trece de la baza zece la orice altă bază B este de a efectua împărţiri succesive ale numărului z prin B; resturile vor fi cifrele numărului în sistemul cu baza B. De exemplu,

	Cît	Rest
109 : 7	15	4
15 : 7	2	1
2 : 7	0	2

109 (în sistemul zecimal) = 214 (în sistemul septimal).

Este natural să ne întrebăm dacă vreo alegere particulară a bazei ar fi cea mai convenabilă. Vom vedea că o bază prea mică prezintă dezavantaje, în timp ce o bază mare necesită învăţarea multor cifre, şi a unei table a înmulţirii foarte mari. Alegerea lui 12 ca bază a fost susţinută datorită faptului că doisprezece se divide exact prin doi, trei, patru şi şase şi, în consecinţă, calculele în care intervin împărţirea şi fracţiile ar fi adeseori simplificate. Pentru a scrie orice întreg în baza doisprezece (sistemul duodecimal), avem nevoie de două noi cifre pentru zece şi unsprezece. Să scriem α în locul lui zece şi β în locul lui unsprezece. Atunci, în sistemul duodecimal ,,doisprezece'' s-ar scrie sub forma 10, ,,douăzeci şi doi'' sub forma 1α, ,,douăzeci şi trei'' sub forma 1β, iar ,,o sută treizeci şi unu'' sub forma $\alpha\beta$.

Inventarea notaţiei poziţionale, atribuită sumerienilor şi babilonenilor şi dezvoltată de hinduşi, a avut o puternică influenţă asupra civilizaţie umane. Primele sisteme de numeraţie erau construite pe principii pur aditive.

Astfel, în sistemul de numerație roman CXVIII reprezintă, „o sută + zece + + cinci + unu + unu + unu”. Sistemele de numerație egiptean, ebraic și grec se aflau la același nivel. Un neajuns al oricărei notații pur aditive este faptul că pe măsură ce numerele cresc sînt necesare din ce în ce mai multe simboluri. Dar principala deficiență a sistemelor antice, ca de pildă cel roman, consta în faptul că însuși procedeul de calcul era foarte dificil; chiar cele mai simple probleme puteau fi rezolvate numai de specialiști. Lucrurile stau cu totul altfel cu sistemul pozițional hindus, care se folosește în prezent. Acesta a fost introdus în Europa medievală de negustorii italieni, care l-au învățat de la musulmani. Sistemul pozițional are avantajul că toate numerele, oricît ar fi de mari sau mici, pot fi scrise cu ajutorul unui număr mic de simboluri (în sistemul zecimal acestea sînt „cifrele arabe” 0, 1, 2, ..., 9). Pe lîngă aceasta, mai avem avantajul important al calculului facil. Regulile de calcul cu numere reprezentate în notație pozițională pot fi rezumate sub forma unor tabele de adunare și înmulțire pentru cifre care pot fi învățate și memorizate o dată pentru totdeauna. Vechea metodă de calcul, rezervată odinioară cîtorva adepți, se predă acum în școala elementară. În istoria culturii se găsesc puține exemple în care progresul științific să fi influențat atît de profund și să fi ușurat atît de mult viața de toate zilele.

3. Calculul în sisteme de numerație diferite de cel zecimal

Folosirea lui zece ca bază a sistemului de numerație se pierde în zorile civilizației și este datorată, fără îndoială, faptului că avem zece degete la mîini, cu care putem număra. Cuvintele pentru desemnarea numerelor în diferite limbi arată însă reminiscențe ale folosirii altor baze, mai cu seamă doisprezece și douăzeci. În engleză și germană cuvintele care servesc la desemnarea lui 11 și 12 nu sînt construite pe principiul zecimal al combinării lui 10 cu cifrele, așa cum sînt „sprezecile”, ci sînt independente din punct de vedere lingvistic de cuvintele folosite pentru desemnarea lui 10. În franceză, cuvintele *vingt* și *quatrevingt* pentru 20 și 80 sugerează că pentru anumite scopuri ar fi putut fi folosit un sistem cu baza 20. În daneză, cuvîntul care desemnează pe 70, *halvfirsindstyve* înseamnă jumătatea drumului de la de trei ori la patrn ori douăzeci. Astronomii Babilonului aveau un sistem de notare care era parțial sexagesimal (cu baza 60), și aceasta pare să explice împărțirea obișnuită a orei și gradului unghiular în 60 de minute.

Într-un sistem diferit de cel zecimal, regulile aritmeticii sînt aceleași, dar tabelele pentru adunarea și înmulțirea acelorași cifre se deosebesc de tabelele din sistemul zecimal. Obișnuiți cu sistemul zecimal și cu limbajul corespunzător prin care desemnăm numerele, s-ar putea ca acest lucru să ni se pară oarecum confuz. Să încercăm cu un exemplu de înmulțire în sistemul septimal. Mai înainte, este indicat să scriem tabele pe care va trebui să le folosim:

Adunarea

	1	2	3	4	5	6
1	2	3	4	5	6	10
2	3	4	5	6	10	11
3	4	5	6	10	11	12
4	5	6	10	11	12	13
5	6	10	11	12	13	14
6	10	11	12	13	14	15

Înmulțirea

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	11	13	15
3	3	6	12	15	21	24
4	4	11	15	22	26	33
5	5	13	21	26	34	42
6	6	15	24	33	42	51

Să înmulțim acum 265 cu 24, unde aceste numere sînt scrise în sistemul septimal. Dacă am scrie numerele în sistemul zecimal atunci ar fi vorba de înmulțirea lui 145 cu 18. Regulile înmulțirii sînt aceleași ca și în sistemul zecimal. Să începem prin a înmulți pe 5 cu 4, ceea ce, după cum se vede din tabela înmulțirii, dă 26,

$$\begin{array}{r} 265 \\ 24 \\ \hline 1456 \\ 563 \\ \hline 10416 \end{array}$$

Scriem pe 6 în locul unităților, apoi trecem pe 2 în locul următor. Mai departe găsim că $4 \cdot 6 = 33$ și că $33 + 2 = 35$. Scriem 5, și continuăm în acest fel pînă cînd efectuăm toate înmulțirile. Adunînd pe $1\,456 + 5\,630$ găsim pe locul unităților $6 + 0 = 6$, iar pe locul șepților pe $5 + 3 = 11$. Scriem 1 și trecem pe 1 pe locul „patruzeci și nouă", unde obținem $1 + 6 + + 4 = 14$. Rezultatul final este $265 \cdot 24 = 10\,416$.

Pentru a verifica rezultatul, înmulțim aceleași numere în sistemul zecimal. Numărul 10 416 (în sistemul septimal) poate fi scris în sistemul zecimal prin găsirea puterilor lui 7 pînă la puterea a patra: $7^2 = 49$, $7^3 = 343$, $7^4 = 2\,401$. De unde rezultă că $10\,416 = 2\,401 + 4 \cdot 49 + 7 + 6$, deoarece termenii din dreapta egalității sînt scriși în sistemul zecimal. Adunînd aceste numere găsim că numărul 10 416 din sistemul septimal este egal cu numărul 2 610 din sistemul zecimal. Să înmulțim acum pe 145 cu 18 în sistemul zeeimal; rezultatul este 2 610, deci calculele corespund.

Exerciții: 1) Alcătuiți tabelele adunării și înmulțirii în sistemul duodecimal și dați cîteva exemple de genul celor de mai sus.

2) Scrieți „treizeci" și „o sută și treizeci și trei" în sistemele cu bazele 5, 7, 11, 12.

3) Ce reprezintă simbolurile 11111 și 21212 în aceste sisteme?

4) Alcătuiți tabelele adunării și înmulțirii pentru sistemele cu bazele 5, 11, 13.

Din punct de vedere teoretic, sistemul pozițional cu baza 2 se evidențiază ca sistemul cu cea mai mică bază posibilă. Singurele cifre în acest *sistem diadic* sînt 0 și 1; orice alt număr z se scrie ca o combinație a acestor simboluri. Tabelele adunării și înmulțirii conțin numai două reguli $1 + 1 = 10$ și $1 \cdot 1 = 1$. Neajunsul acestui sistem este însă evident: pentru a reprezenta numere mici sînt necesare expresii lungi. Astfel, numărul șaptezeci și nouă, care se exprimă sub forma $1 \cdot 2^6 + 0 \cdot 2^5 + 0 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2 + 1$, se scrie în sistemul diadic ca 1001111.

Pentru a ilustra cît de simplă este înmulțirea în sistemul diadic, vom înmulți numerele șapte cu cinci, care se scriu respectiv sub forma 111 și 101. Avînd în vedere că în acest sistem $1 + 1 = 10$, avem

$$\begin{array}{r} 111 \\ 101 \\ \hline 111 \\ 111 \\ \hline 100011 \end{array} = 2^5 + 2 + 1,$$

prin urmare, cum era și de așteptat, treizeci și cinci.

Gottfried Wilhelm Leibniz (1646—1716), una dintre cele mai luminate minți ale timpului său, aprecia foarte mult sistemul diadic. Iată ce spunea cu privire la acesta Laplace: „Leibniz vedea în aritmetica sa binară imaginea creației. Lui i se părea că unitatea reprezenta pe Dumnezeu, iar zero — vidul, și că Ființa Supremă a creat toate ființele din neant, exact în același fel cum unitatea și zero exprimă toate numerele în sistemul lui de numerație".

Exercițiu: Considerați problema reprezentării întregilor în sistemul cu baza a. Pentru a numi numerele în acest sistem, avem nevoie de denumiri pentru cifrele $0, 1, \ldots, a-1$ și pentru diferitele puteri ale lui a: $a, a^2, a^3, \ldots$. Cîte denumiri pentru numere sînt necesare pentru a numi toate numerele de la zero la o mie, dacă $a = 2, 3, 4, 5, \ldots, 15$? Care bază pretinde cel mai mic număr de denumiri? (Exemple: dacă $a = 10$, avem nevoie de zece denumiri pentru cifre, apoi încă trei denumiri pentru 10, 100 și 1000, în total 13. Pentru $a = 20$, avem nevoie de douăzeci de cuvinte pentru desemnarea cifrelor, pe lîngă cuvintele pentru 20 și 400, deci în total 22. Dacă $a = 100$, ajungem la 101.)

•§ 2. INFINITUDINEA SISTEMULUI DE NUMERE. INDUCȚIA MATEMATICĂ

1. Principiul inducției matematice

Șirul numerelor întregi 1, 2, 3, 4, ... este nesfîrșit, pentru că după orice întreg n putem scrie întregul următor, $n + 1$. Exprimăm această proprietate a șirului de numere întregi spunînd că există o *infinitate* de întregi. Șirul

întregilor reprezintă cel mai simplu şi mai natural exemplu al infinitului matematic, care joacă un rol dominant în matematica modernă. În nenumărate locuri din această carte vom avea de-a face cu colecţii sau „mulţimi", care conţin o infinitate de obiecte matematice, cum ar fi, de exemplu, mulţimea tuturor punctelor de pe o dreaptă, sau mulţimea tuturor triunghiurilor din plan. Dar, şirul infinit al întregilor este, fără îndoială, cel mai simplu exemplu de mulţime infinită.

Procedeul treptat de trecere de la n la $n + 1$, care generează şirul infinit al întregilor, formează împreună cu acesta baza unuia din cele mai importante modele de raţionament matematic, principiul inducţiei matematice. „Inducţia empirică" aplicată în ştiinţele naturii, pornind de la un şir particular de observaţii asupra unui anumit fenomen, ajunge la enunţarea unei legi generale care guvernează toate manifestările acestui fenomen. Gradul de certitudine cu care este stabilită legea depinde de numărul observaţiilor individuale şi concluziile deduse din ele. Adesea acest gen de raţionament inductiv este pe deplin convingător; propoziţia că Soarele va răsări mîine la est este tot atît de sigură cît poate fi în general vreun lucru; cu toate acestea, caracterul acestei afirmaţii este cu totul altul decît al unei teoreme, demonstrate printr-un raţionament strict logic sau matematic.

Într-un mod cu totul diferit, *inducţia matematică* este folosită pentru stabilirea adevărului unei teoreme matematice într-un şir infinit de cazuri, primul, al doilea, al treilea şi aşa mai departe, fără nici o excepţie. Să notăm cu A o propoziţie oarecare care se referă la un întreg arbitrar n. De exemplu, A poate fi propoziţia „suma unghiurilor într-un poligon convex cu $n + 2$ laturi este egală cu $n \cdot 180°$". Sau, A' poate fi afirmaţia: „trasînd n drepte într-un plan nu putem împărţi planul în mai mult de 2^n părţi". Pentru a demonstra o astfel de teoremă pentru *orice* întreg n nu este suficient să o demonstrăm pentru primele 10, sau 100, sau chiar 1000 de valori ale lui n. Aceasta ar corespunde de fapt inducţiei empirice. În schimb, trebuie să folosim un raţionament riguros matematic şi neempiric, al cărui caracter va reieşi din demonstrarea exemplelor particulare date A şi A'. În cazul A, ştim că pentru $n = 1$ poligonul este un triunghi, şi din geometria elementară ştim că suma unghiurilor este egală, în acest caz, cu $1 \cdot 180°$. În cazul unui patrulater, $n = 2$, trasăm o diagonală care împarte patrulaterul în două triunghiuri, şi de data aceasta se vede imediat că suma unghiurilor patrulaterului este egală cu suma unghiurilor celor două triunghiuri, ceea ce dă $180° + 180° = 2 \cdot 180°$. Trecînd la cazul unui poligon cu cinci laturi, $n = 3$, îl descompunem într-un triunghi şi un patrulater. Deoarece ultimul are suma unghiurilor egală cu $2 \cdot 180°$, după cum am arătat, şi deoarece triunghiul are suma unghiurilor egală cu $180°$, obţinem $3 \cdot 180°$ pentru pentagon. Acum este clar că putem continua raţionamentul indefinit pe aceeaşi cale, demonstrînd teorema pentru $n = 4$, apoi pentru $n = 5$ şi aşa mai departe.

Orice propoziție se deduce, în același mod, din cea precedentă, astfel încît teorema generală poate fi stabilită pentru orice n.

În mod similar, putem demonstra teorema A'. Pentru $n = 1$ ea este evident adevărată, deoarece orice dreaptă împarte planul în două părți. Să adăugăm a doua dreaptă. Fiecare din părțile precedente va fi împărțită în două noi părți, afară de cazul în care noua dreaptă este paralelă cu prima. În orice caz, pentru $n = 2$, avem cel mult $4 = 2^2$ părți. Să adăugăm a treia dreaptă. Fiecare din părțile precedente va fi împărțită în două părți, sau va rămîne neschimbată. Astfel, numărul părților nu este mai mare decît $2^2 \cdot 2 = 2^3$. Știind că acest lucru este adevărat, putem trata cazul următor în același mod și așa mai departe.

Ideea esențială a raționamentelor precedente este de a stabili o teoremă generală A pentru orice valoare a lui n, demonstrînd succesiv această teoremă pentru un șir infinit de cazuri particulare $A_1, A_2, \ldots$. Posibilitatea acestui raționament se bazează pe următoarele premise: a) există o metodă generală pentru a arăta că *dacă* o propoziție A_r este adevărată, atunci *și* propoziția următoare A_{r+1} va fi adevărată; b) prima propoziție A_1 este *cu siguranță* adevărată. Faptul că aceste două condiții sînt suficiente pentru stabilirea adevărului *tuturor* propozițiilor $A_1, A_2, A_3, \ldots$ reprezintă un principiu logic, tot atît de fundamental pentru matematică, ca și regulile clasice ale logicii lui Aristotel. Formulăm acest principiu în modul următor:

Să presupunem că vrem să stabilim un șir infinit de propoziții matematice

$$A_1, A_2, A_3, \ldots$$

care, la un loc, formează propoziția generală A. *Să presupunem că* a) *printr-un raționament matematic s-a arătat că dacă r este un întreg și dacă despre afirmația A_r știm că este adevărată*, atunci *adevărul afirmației A_{r+1} rezultă*, și că b) știm *despre propoziția A_1 că este adevărată. Atunci toate propozițiile șirului nostru trebuie să fie adevărate și, prin urmare, A este demonstrată.*

Vom accepta fără ezitare acest principiu (tot așa cum acceptăm toate regulile logicii obișnuite) ca un principiu de bază al raționamentului matematic. Într-adevăr putem stabili adevărul oricărei propoziții A_n, pornind de la propoziția dată b), anume că A_1 este adevărată, și continuînd, prin folosirea repetată a afirmației a), pentru a stabili succesiv adevărul lui $A_2, A_3, A_4, \ldots$ etc., pînă ce ajungem la propoziția A_n. Principiul inducției matematice se bazează în acest mod pe faptul că după orice întreg r urmează un alt întreg, $r + 1$ și că, pornind de la întregul 1, putem ajunge, după un număr finit de astfel de pași, la orice întreg n.

Adesea, principiul inducției matematice se aplică fără a-l menționa explicit, sau este indicat pur și simplu prin obișnuitul „etc.” sau „și așa mai departe”. Această formă se întrebuințează în special în predarea matematicii

elementare. În demonstrarea unor teoreme mai subtile este însă inevitabilă folosirea explicită a acestui principiu. Vom da în cele ce urmează cîteva exemple simple, dar nu chiar banale.

2. Progresia aritmetică

Pentru orice valoare a lui n, *suma* $1 + 2 + 3 + \ldots + n$ *a primilor* n *întregi este egală cu* $\frac{n(n+1)}{2}$. Pentru a demonstra această teoremă prin inducție matematică, trebuie să arătăm că pentru orice n, afirmația A_n:

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2} \tag{1}$$

este adevărată. a) Observăm că dacă r este un întreg și dacă propoziția A_r este adevărată, adică dacă știm că

$$1 + 2 + 3 + \ldots + r = \frac{r(r+1)}{2},$$

atunci adunînd numărul $(r + 1)$ ambilor membri ai acestei egalități, obținem egalitatea

$$1 + 2 + 3 + \ldots + r + (r+1) = \frac{r(r+1)}{2} + (r+1) =$$

$$= \frac{r(r+1) + 2(r+1)}{2} = \frac{(r+1)(r+2)}{2},$$

care este chiar propoziția A_{r+1}. b) Propoziția A_1 este evident adevărată, deoarece $1 = \frac{1 \cdot 2}{2}$. Deci, conform principiului inducției matematice, propoziția A_n este adevărată pentru orice n, ceea ce trebuia să demonstrăm.

De obicei, această teoremă se demonstrează scriind suma $1 + 2 + 3 + \ldots + n$, sub două forme:

$$S_n = 1 + 2 + \ldots + (n-1) + n$$

și

$$S_n = n + (n-1) + \ldots + 2 + 1.$$

Adunînd, vedem că fiecare pereche de numere, aflate în aceeași coloană, dă ca sumă pe $n + 1$ și, deoarece în total sînt n coloane, rezultă că

$$2S_n = n(n+1),$$

ceea ce demonstrează rezultatul dorit.

Din (1) deducem imediat formula sumei primilor $n + 1$ termeni ai unei *progresii aritmetice*,

$$(2) \qquad P_n = a + (a + d) + (a + 2d) + \ldots + (a + nd) = \\ = \frac{(n + 1)(2a + nd)}{2}.$$

Într-adevăr

$$P_n = (n + 1)a + (1 + 2 + \ldots + n)\, d = (n + 1)a + \frac{n(n + 1)d}{2} =$$

$$= \frac{2(n + 1)a + n(n + 1)d}{2} = \frac{(n + 1)(2a + nd)}{2}.$$

În cazul $a = 0$, $d = 1$, ultima egalitate devine (1).

3. Progresia geometrică

Putem trata în acelaşi mod progresia geometrică generală. Vom demonstra că pentru orice valoare a lui n avem

$$(3) \qquad G_n = a + aq + aq^2 + \ldots + a\, q^n = a\frac{1 - q^{n+1}}{1 - q}.$$

(Să presupunem că $q \neq 1$, deoarece în caz contrar, membrul drept din (3) nu are sens.)

Deci, în mod sigur, afirmaţia noastră este adevărată pentru $n = 1$, deoarece în acest caz avem

$$G_1 = a + a\, q = \frac{a(1 - q^2)}{1 - q} = \frac{a(1 + q)(1 - q)}{1 - q} = a(1 + q).$$

Şi *dacă* presupunem că

$$G_r = a + aq + \ldots + aq^r = a\frac{1 - q^{r+1}}{1 - q},$$

atunci găsim

$$G_{r+1} = (a + aq + \dots + aq^r) + aq^{r+1} = G_r + aq^{r+1} = a\frac{1 - q^{r+1}}{1 - q} +$$

$$+ aq^{r+1} = \frac{a(1 - q^{r+1}) + q^{r+1}(1 - q)}{1 - q} =$$

$$= a\frac{1 - q^{r+1} + q^{r+1} - q^{r+2}}{1 - q} = a\frac{1 - q^{r+2}}{1 - q}.$$

Aceasta este însă chiar afirmația (3) pentru $n = r + 1$. Demonstrația este prin urmare completă.

În manualele elementare, se dă altă demonstrație. Să punem

$$G_n = a + aq + \dots + aq^n,$$

înmulțind ambii membri ai acestei egalități cu q, obținem

$$qG_n = aq + aq^2 + \dots + aq^{n+1}.$$

Scăzînd după aceea membru cu membru cele două egalități vom obține:

$$G_n - qG_n = a - aq^{n+1},$$

$$(1 - q)G_n = a(1 - q^{n+1}),$$

$$G_n = a\frac{1 - q^{n+1}}{1 - q}.$$

4. Suma primelor n pătrate

Următoarea aplicație interesantă a principiului inducției matematice se referă la suma primelor n pătrate. Prin încercări directe găsim că cel puțin pentru valori mici ale lui n avem

$$(4) \qquad 1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n + 1)(2n + 1)}{6}$$

și am putea *presupune* că această formulă remarcabilă este valabilă pentru orice n întreg pozitiv. Pentru a *demonstra* acest lucru, vom folosi din nou principiul inducției matematice. Observăm mai întîi că dacă afirmația A_n, care în acest caz este egalitatea (4), este adevărată pentru $n = r$, astfel încît

$$1^2 + 2^2 + 3^2 + \dots + r^2 = \frac{r(r + 1)(2r + 1)}{6},$$

atunci, adunînd $(r+1)^2$ ambilor membri ai acestei egalități, obținem

$$1^2+2^2+3^2+\ldots+r^2+(r+1)^2=\frac{r(r+1)(2r+1)}{6}+(r+1)^2=$$

$$=\frac{r(r+1)(2r+1)+6(r+1)^2}{6}=\frac{(r+1)[r(2r+1)+6(r+1)]}{6}=$$

$$=\frac{(r+1)(2r^2+7r+6)}{6}=\frac{(r+1)(r+2)(2r+3)}{6},$$

care este tocmai afirmația A_{r+1}, deoarece ea se obține din relația (4) prin înlocuirea lui r cu $r+1$. Pentru a completa demonstrația, trebuie să mai remarcăm numai că afirmația A_1, care se reduce la egalitatea

$$1^2=\frac{1(1+1)(2+1)}{6}$$

este adevărată. Deci egalitatea (4) este adevărată pentru orice n.

Formule de același gen pot fi găsite pentru puteri mai mari ale întregilor, $1^k+2^k+3^k+\ldots+n^k$, unde k este un întreg pozitiv oarecare. Ca exercițiu, cititorul poate demonstra cu ajutorul inducției matematice formula

$$(5)\qquad 1^3+2^3+3^3+\ldots+n^3=\left[\frac{n(n+1)}{2}\right]^2.$$

Trebuie neapărat să remarcăm că deși principiul inducției matematice este suficient pentru a *demonstra* formula (5), dacă această formulă a fost deja obținută, totuși demonstrația nu dă nici o indicație referitoare la modul în care s-a ajuns la această formulă; de ce trebuie să ghicim că tocmai expresia $[n(n+1)/2]^2$ dă suma primilor n cuburi, și s-o preferăm pe aceasta altor expresii de aceeași formă $[n(n+1)/3]^2$, sau $(19n^2-41n+24)/2$. Faptul că demonstrația unei teoreme constă în aplicarea anumitor reguli simple de logică nu ne scutește de elementul creator în matematici, care se manifestă prin efectuarea unei alegeri dintr-o infinitate de cazuri posibile. Problema originii *ipotezei* (5) aparține unui domeniu în care nu există nici un fel de reguli generale; experiența, analogia și inducția constructivă își au contribuția lor. Dar o dată ce ipoteza corectă este formulată, principiul inducției matematice este adesea suficient pentru a furniza demonstrația. În măsura în care o astfel de demonstrație nu dă un indiciu asupra actului descoperirii, ar putea fi numită, mai potrivit, o *verificare*.

*5. O inegalitate importantă

Într-unul din capitolele următoare vom folosi inegalitatea

$$(6) \qquad (1 + p)^n \geqslant 1 + np,$$

care este valabilă pentru orice număr $p > -1$ și orice întreg pozitiv n. (Pentru generalitate, anticipăm aici folosirea numerelor negative și neîntregi, presupunînd că p poate fi orice număr mai mare decît -1. Demonstrația pentru cazul general este exact aceeași ca și în cazul în care p este un întreg pozitiv). Vom folosi și de data aceasta inducția matematică.

a) *Dacă* este adevărat că $(1 + p)^r \geqslant 1 + rp$, atunci înmulțind ambii membri ai acestei inegalități cu numărul pozitiv $1 + p$, obținem

$$(1 + p)^{r+1} \geqslant 1 + rp + p + rp^2.$$

Renunțînd la termenul pozitiv rp^2 nu facem decît să întărim această inegalitate, astfel încît

$$(1 + p)^{r+1} \geqslant 1 + (r + 1)p,$$

ceea ce arată că inegalitatea (6) rămîne valabilă și pentru întregul următor, $r + 1$. b) Este evident adevărat că $(1 + p)^1 \geqslant 1 + p$. Aceasta completează demonstrația faptului că (6) este adevărată pentru orice n. Limitarea la numerele $p > -1$ este esențială. Dacă $p < -1$, atunci $1 + p$ este negativ și raționamentul din a) își pierde valabilitatea, deoarece dacă ambii membri ai primei inegalități sînt înmulțiți cu o cantitate negativă, inegalitatea își schimbă sensul (de exemplu, dacă înmulțim ambii membri ai inegalității $3 > 2$ cu -1, obținem $-3 > -2$, ceea ce nu este adevărat).

6. Formula binomului

Adesea este important să avem o expresie explicită pentru puterea a n-a a unui binom, $(a + b)^n$. Prin calcul explicit găsim că

pentru $n = 1$,

$$(a + b)^1 = a + b,$$

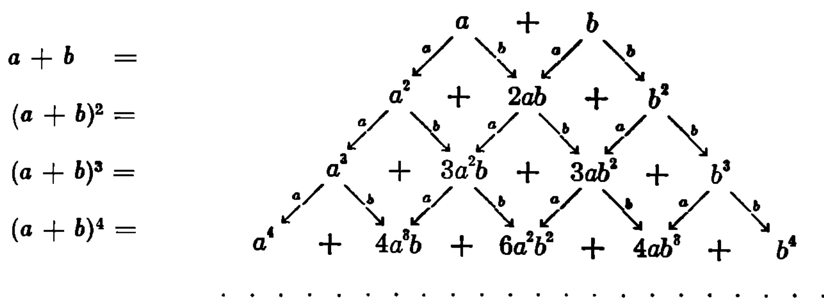
pentru $n = 2$,

$$(a + b)^2 = (a + b)(a + b) = a(a + b) + b(a + b) =$$
$$= a^2 + 2ab + b^2,$$

pentru $n = 3$,

$$(a + b)^3 = (a + b)(a + b)^2 = a(a^2 + 2ab + b^2) + $$
$$+ b(a^2 + 2ab + b^2) = a^3 + 3a^2b + 3ab^2 + b^3,$$

şi aşa mai departe. Care este legea generală de formare, care se află în spatele cuvintelor „şi aşa mai departe"? Să examinăm procesul prin care am calculat pe $(a + b)^2$. Deoarece $(a + b)^2 = (a + b)(a + b)$, obţinem expresia lui $(a + b)^2$ înmulţind fiecare termen din expresia lui $a + b$ cu a, apoi cu b şi adunînd. Acelaşi procedeu a fost folosit pentru a calcula pe $(a + b)^3 = (a + b)(a + b)^2$. Putem continua în acelaşi mod, pentru a calcula $(a + b)^4$, $(a + b)^5$ şi aşa mai departe. Expresia lui $(a + b)^n$ va fi obţinută prin înmulţirea fiecărui termen al expresiei obţinute anterior pentru $(a + b)^{n-1}$ mai întîi cu a, apoi cu b, şi adunînd apoi rezultatul. Aceasta duce la următoarea diagramă:



ceea ce dă dintr-o dată regula generală de formare a coeficienţilor din expresia lui $(a + b)^n$. Să construim un tabel triunghiular din numere naturale, începînd cu coeficienţii 1, 1 ai lui $a + b$, astfel încît fiecare număr al triunghiului să fie suma celor două numere aflate de o parte şi de alta a lui, în linia precedentă. Acest tabel este cunoscut sub numele de *Triunghiul lui Pascal.*

1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1

A n-*a linie a acestui tabel dă coeficienții din dezvoltarea lui* (a + b)ⁿ, *după duterile descrescătoare ale lui* a *și puterile crescătoare ale lui* b ; astfel

$$(a + b)^7 = a^7 + 7a^6b + 21a^5b^2 + 35a^4b^3 + 35a^3b^4 + 21a^2b^5 + 7ab^6 + b^7.$$

Folosind o notație concisă, cu indici, putem nota numerele aflate în cea de a n-a linie a triunghiului lui Pascal sub forma

$$C_n^0 = 1,\ C_n^1,\ C_n^2,\ C_n^3,\ \ldots,\ C_n^{n-1},\ C_n^n = 1.$$

Atunci, formula generală pentru $(a + b)^n$ poate fi scrisă

$$(7)\qquad (a + b)^n = a^n + C_n^1 a^{n-1}b + C_n^2 a^{n-2}b^2 + \ldots + C_n^{n-1}\, ab^{n-1} + b^n.$$

Conform legii de formare a triunghiului lui Pascal avem relația

$$(8)\qquad C_n^i = C_{n-1}^{i-1} + C_{n-1}^i.$$

Ca exercițiu, cititorul oarecum experimentat poate folosi această relație împreună cu faptul că $C_1^0 = C_1^1 = 1$, pentru a arăta prin inducție matematică că

$$(9)\qquad C_n^i = \frac{n(n-1)(n-2)\ldots(n-i+1)}{1\cdot 2\cdot 3\ldots i} = \frac{n!}{i!(n-i)!}.$$

Pentru orice întreg pozitiv n, simbolul $n!$ (se citește „factorial de n") reprezintă produsul primilor n întregi: $n! = 1\cdot 2\cdot 3\ldots n$. De asemenea, este comod să notăm $0! = 1$, astfel încît (9) este valabilă și pentru $i = 0$ și $i = n$. Această formulă explicită pentru coeficienții dezvoltării binomului se numește uneori *formula binomului*. (A se vedea de asemenea p.495).

Exerciții: Demonstrați prin inducție matematică următoarele egalități:

1) $\dfrac{1}{1\cdot 2} + \dfrac{1}{2\cdot 3} + \ldots + \dfrac{1}{n(n+1)} = \dfrac{n}{n+1}$.

2) $\dfrac{1}{2} + \dfrac{2}{2^2} + \dfrac{3}{2^3} + \ldots + \dfrac{n}{2^n} = 2 - \dfrac{n+2}{2^n}$.

*3) $1 + 2q + 3q^2 + \ldots + nq^{n-1} = \dfrac{1-(n+1)q^n + nq^{n+1}}{(1-q)^2}$.

*4) $(1+q)(1+q^2)(1+q^4)\ldots(1+q^{2^n}) = \dfrac{1-q^{2^{n+1}}}{1-q}$.

Găsiți suma următoarelor progresii geometrice:

5) $\dfrac{1}{1+x^2} + \dfrac{1}{(1+x^2)^2} + \ldots + \dfrac{1}{(1+x^2)^n}$.

6) $1 + \frac{x}{1+x^2} + \frac{x^2}{(1+x^2)^2} + \ldots + \frac{x^n}{(1+x^2)^n}$.

7) $\frac{x^2-y^2}{x^2+y^2} + \left(\frac{x^2-y^2}{x^2+y^2}\right)^2 + \ldots + \left(\frac{x^2-y^2}{x^2+y^2}\right)^n$.

Folosind formulele (4) şi (5) demonstraţi că:

*8) $1^2 + 3^2 + \ldots + (2n+1)^2 = \frac{(n+1)(2n+1)(2n+3)}{3}$.

*9) $1^3 + 3^3 + \ldots + (2n+1)^3 = (n+1)^2(2n^2+4n+1)$.

10) Demonstraţi aceleaşi rezultate cu ajutorul inducţiei matematice.

*7. Alte observaţii asupra inducţiei matematice

Principiul inducţiei matematice poate fi puţin generalizat: „Dacă un şir de propoziţii A_s, $A_{s+1}, A_{s+2}, \ldots$ este dat, unde s este un întreg pozitiv şi dacă

a) pentru orice valoare a lui $r \geq s$, valabilitatea lui A_{r+1} rezultă din valabilitatea lui A_r, şi

b) se ştie că A_s este adevărată,

atunci toate propoziţiile A_s, A_{s+1}, A_{s+2} sînt adevărate; adică, A_n este adevărată pentru orice $n \geq s$". Exact acelaşi raţionament, folosit pentru stabilirea valabilităţii principiului obişnuit al inducţiei matematice se aplică şi aici, înlocuind şirul 1, 2, 3, ... prin şirul asemănător d, $s+1$, $s+2$, $s+3$, ... Folosind principiul sub această formă, putem întări puţin inegalitatea de la p. 32, eliminînd posibilitatea semnului „=". Afirmăm că: pentru orice $\neq 0$ şi $p > -1$ şi *orice întreg* $n \geq 2$ *avem*

(10) $$(1+p)^n > 1 + np.$$

Demonstraţia va fi lăsată pe seama cititorului.

Strîns legat de principiul inducţiei matematice este „principiul celui mai mic număr întreg" care afirmă că *orice mulţime nevidă* C, *de întregi pozitivi, are un cel mai mic element.* O mulţime este vidă dacă nu are elemente, de exemplu, mulţimea cercurilor drepte sau mulţimea întregilor n, astfel încît $n > n$. Datorită unor motive evidente vom exclude astfel de mulţimi în enunţul principiului. Mulţimea C poate fi finită, ca de pildă mulţimea 1, 2, 3, 4, 5, sau infinită, ca mulţimea tuturor numerelor pare 2, 4, 6, 8, 10, ... Orice mulţime nevidă C trebuie să conţină cel puţin un număr întreg, de pildă n, iar cel mai mic dintre întregii 1, 2, 3, ..., n, care aparţin lui C, va fi cel mai mic număr întreg din C.

Singurul mod în care putem înţelege importanţa acestui principiu este de a observa că el *nu* este valabil în cazul mulţimilor C, care nu sînt formate numai din numere întregi; de exemplu, mulţimea fracţiilor pozitive 1, 1/2, 1/3, 1/4, ... nu conţine un cel mai mic element.

Din punct de vedere pur logic, este interesant de observat că, cu ajutorul principiului celui mai mic număr întreg, se poate *demonstra* principiul inducţiei matematice ca teoremă. Într-adevăr, să considerăm un şir de propoziţii A_1, A_2, A_3, ... astfel încît

a) pentru orice întreg pozitiv r, valabilitatea lui A_{r+1} rezultă din aceea a lui A_r;

b) se ştie că A_1 este adevărată.

Vom demonstra că ipoteza „că una din propozițiile A este falsă" duce la contradicție. Într-adevăr, dacă măcar una din propozițiile A ar fi falsă, mulțimea C a tuturor întregilor pozitivi n pentru care A_n este falsă ar fi nevidă. Atunci, conform principiului celui mai mic număr întreg, C ar conține un cel mai mic întreg p, care trebuie să fie > 1, datorită lui b). Dar atunci A_p ar fi falsă, iar A_{p-1} adevărată. Aceasta contrazice ipoteza a).

Subliniem încă o dată că principiul inducției matematice se deosebește esențial de inducția empirică din științele naturii. Confirmarea unei legi generale într-un număr finit de cazuri, oricît de mare ar fi acest număr, nu constituie o demonstrație în sensul matematic riguros al cuvîntului, chiar dacă pentru moment nu se cunoaște nici o singură excepție. În astfel de împrejurări afirmația sau „legea" ar rămîne doar o *ipoteză* foarte rezonabilă, supusă modificării datorită rezultatelor experimentelor viitoare. În matematică, o lege sau o teoremă se consideră demonstrată numai dacă se poate arăta că este o consecință logică, necesară, a anumitor ipoteze care sînt acceptate ca valabile. Există multe exemple de propoziții matematice care au fost verificate în fiecare caz particular considerat pînă acum, dar pentru care nu s-a găsit încă o demonstrație generală (a se vedea, de exemplu, p. 46). Putem *bănui* că o teoremă este în general adevărată dacă se dovedește a fi valabilă într-un număr mare de cazuri, și atunci putem încerca *s-o demonstrăm* prin inducție matematică. Dacă încercarea reușește, teorema este demonstrată; în caz contrar, teorema poate fi adevărată sau falsă și poate fi demonstrată sau infirmată mai devreme sau mai tîrziu, prin alte metode.

Folosind principiul inducției matematice, trebuie să fim întotdeauna siguri că condițiile a) și b) sînt într-adevăr satisfăcute. Neglijarea acestei precauții poate duce la absurdități, ca de pildă următoarea, în care cititorul este invitat să descopere greșeala. Vom „demonstra" că *oricare două numere pozitive sînt egale*; de exemplu, că $5 = 10$.

Începem cu o definiție: dacă a și b sînt două numere întregi pozitive diferite, definim $\max(a,b)$ ca fiind cel mai mare dintre numerele a sau b; dacă $a = b$ punem $\max(a, b) = a = b$. Astfel, $\max(3,5) = \max(5,3) = 5$, în timp ce $\max(4,4) = 4$. Acum, fie A_n propoziția „dacă a și b sînt două numere întregi pozitive astfel încît $\max(a, b) = n$, atunci $a = b$".

a) Să presupunem că A_r este adevărată. Fie a și b două numere întregi pozitive, astfel încît $\max(a,b) = r + 1$. Să considerăm întregii

$$\alpha = a - 1$$
$$\beta = b - 1;$$

atunci $\max(\alpha, \beta) = r$. Deci $\alpha = \beta$, deoarece presupunem că A_r este adevărată. Rezultă că $a = b$; deci A_{r+1} este adevărată.

b) A_1 este evident adevărată, deoarece dacă $\max(a,b) = 1$, atunci a și b (prin ipoteză numere întregi pozitive) trebuie să fie, fiecare egale cu 1. Astfel, conform principiului inducției matematice, A_n este adevărată pentru orice n.

Fie acum a și b două numere întregi pozitive oarecare; să notăm $\max(a, b) = r$, Deoarece, după cum s-a arătat, A_n este adevărată pentru orice n, rezultă în particular că A_r este adevărată. Deci, $a = b$.

TEORIA NUMERELOR

INTRODUCERE

Reprezentările mistice şi superstiţiile legate de numerele întregi s-au stins încetul cu încetul, dar interesul matematicienilor pentru aceste numere nu a slăbit niciodată. Euclid (aproximativ 300 î.e.n.), a cărui faimă este legată îndeosebi de acea parte a *Elementelor* sale dedicată bazelor geometriei, studiată în prezent în şcoală, pare să fi adus contribuţii originale în domeniul teoriei numerelor, în timp ce geometria lui a fost în mare parte o compilaţie a rezultatelor obţinute anterior. Diofant din Alexandria (aproximativ 275 e.n.), unul din primii algebrişti, şi-a lăsat de asemenea amprenta în teoria numerelor. Pierre de Fermat (1601—1665) din Toulouse, jurist de profesie, unul dintre cei mai mari matematicieni ai timpului său, a pus bazele teoriei moderne în acest domeniu. Euler (1707—1783), cel mai fecund matematician, a obţinut în cercetările sale rezultate apreciabile referitoare la teoria numerelor. Nume proeminente în analele matematicii, ca Legendre, Dirichlet, Riemann, trebuie adăugate acestei liste. Gauss (1777—1855), cel mai mare matematician al timpurilor moderne, care s-a devotat multor ramuri ale matematicii, vorbind despre teoria numerelor, a spus: „Matematica este regina ştiinţelor, iar teoria numerelor este regina matematicii".

§ 1. NUMERELE PRIME

1. Rezultate fundamentale

Cele mai multe propoziţii din teoria numerelor, ca şi din matematică în general, se referă nu la un singur obiect — la numărul 5 sau la numărul 32 — ci la o întreagă clasă de obiecte, care au o proprietate comună, ca de pildă clasa tuturor întregilor pari

$$2, 4, 6, 8, \ldots$$

sau clasa tuturor întregilor divizibili cu 3,

$$3, 6, 9, 12, \ldots$$

sau clasa tuturor pătratelor de întregi,

$$1, 4, 9, 16, \ldots$$

şi aşa mai departe.

De importanţă fundamentală în teoria numerelor este clasa tuturor *numerelor prime*. Cei mai mulţi întregi pot fi descompuşi în factori mai mici: $10 = 2 \cdot 5$, $111 = 3 \cdot 37$, $144 = 3 \cdot 3 \cdot 2 \cdot 2 \cdot 2 \cdot 2$ etc. Numerele care nu pot fi descompuse în acest mod se numesc numere prime. Mai precis, *un număr prim este un întreg* p, *mai mare decît unu, care nu are alţi factori în afară de el însuşi şi de unu.* (Un întreg a se numeşte *factor* sau *divizor* al unui întreg b, dacă există un întreg c, astfel încît $b = ac$). Numerele 2, 3, 5, 7, 11, 13, 17, ... sînt prime, în timp ce 12, de exemplu, nu este prim, pentru că $12 = 3 \cdot 4$. Importanţa clasei numerelor prime constă în faptul că *orice* întreg poate fi exprimat ca *produs de numere prime:* dacă un număr nu este prim, el poate fi factorizat succesiv pînă ce toţi factorii sînt primi; astfel, de exemplu,

$$360 = 3 \cdot 120 = 3 \cdot 30 \cdot 4 = 3 \cdot 3 \cdot 10 \cdot 2 \cdot 2 =$$
$$= 3 \cdot 3 \cdot 5 \cdot 2 \cdot 2 \cdot 2 = 2^3 \cdot 3^2 \cdot 5.$$

Un întreg (diferit de 0 sau 1) care nu este prim se numeşte *număr compus*.

Una din primele întrebări care se pun, referitoare la clasa numerelor prime, este dacă există numai un număr finit de numere prime diferite sau dacă clasa numerelor prime conţine o infinitate de elemente, ca şi clasa tuturor întregilor, a cărei parte este. Răspunsul este următorul: *există o infinitate de numere prime.*

Demonstraţia faptului că clasa numerelor prime este infinită, aşa cum a fost dată de Euclid, reprezintă un model de raţionament matematic. La baza ei stă „metoda indirectă". Începem prin a încerca să presupunem că teorema este falsă. Aceasta înseamnă că ar exista numai un număr finit de numere prime, eventual foarte multe — un milion sau aşa ceva — sau, folosind o notaţie foarte generală şi nerestrictivă, să presupunem că acest număr este n. Folosind notaţia indicială, putem nota toate aceste numere prime cu $p_1, p_2, \ldots, p_n$. Orice alt număr va fi compus şi trebuie să fie divizibil cel puţin prin unul din numerele prime $p_1, p_2, p_3, \ldots, p_n$. Ajungem deci la o contradicţie, construind un număr A, care diferă de fiecare din numerele prime $p_1, p_2, \ldots, p_n$, deoarece este mai mare decît fiecare din ele, şi care nu se divide cu nici unul din ele. Acest număr este

$$A = p_1 p_2 \ldots p_n + 1,$$

adică 1 plus produsul tuturor numerelor prime presupuse. Numărul A este mai mare decît oricare din numerele p şi deci trebuie să fie compus. Dar A împărţit cu p_1 sau cu p_2 etc. dă întotdeauna restul 1; de aceea, A nu se divide cu nici unul din numerele p. Deoarece ipoteza noastră iniţială că există doar un număr finit de numere prime duce la o contradicţie, deducem că ipoteza este absurdă şi deci contrariul ei trebuie să fie adevărat. Aşadar, teorema este demonstrată.

Cu toate că această demonstraţie este indirectă, ea poate fi modificată cu uşurinţă pentru a da o metodă de a construi, cel puţin teoretic, un şir infinit de numere prime. Să presupunem că pornind de la un număr prim oarecare, să spunem $p_1 = 2$, am găsit deja n numere prime $p_1, p_2, \ldots, p_n$; observăm atunci că numărul $p_1 p_2 \ldots p_n + 1$ este fie un număr prim, fie conţine ca factor un număr prim, diferit de numerele deja găsite. Deoarece un astfel de factor poate fi găsit întotdeauna prin încercări directe, sîntem siguri că vom găsi, în orice caz, cel puţin un număr prim p_{n+1}; procedînd în continuare în acest mod, constatăm că şirul numerelor prime, pe care îl putem într-adevăr construi, nu are sfîrşit.

Exerciţiu. Efectuaţi această construcţie începînd cu $p_1 = 2$, $p_2 = 3$ şi găsiţi încă cinci numere prime.

Dacă un număr oarecare poate fi exprimat sub forma unui produs de factori primi, atunci putem aranja aceşti factori în orice ordine. Făcînd abstracţie de ordinea factorilor, ajungem foarte repede la concluzia că descompunerea unui număr N în factori primi este unică: *orice întreg* N *mai mare decît* 1 *poate fi descompus într-un singur mod în produs de numere prime.* Această propoziţie pare la prima vedere atît de evidentă, încît nespecialistul este de obicei înclinat să respingă necesitatea demonstrării ei. Propoziţia amintită nu este nicidecum o banalitate, iar demonstraţia, deşi foarte elementară, necesită un raţionament subtil. Demonstraţia clasică dată de Euclid acestei „teoreme fundamentale a aritmeticii" se bazează pe metoda („algoritmul") găsirii celui mai mare divizor comun a două numere. Despre această metodă vom discuta însă la p. 58. Aici vom da, în schimb, o demonstraţie mai recentă, ceva mai scurtă şi poate mai artificială decît cea a lui Euclid. Ea reprezintă, de asemenea, un exemplu tipic de demonstraţie indirectă. Vom admite existenţa unui întreg care poate fi descompus sub forma unui produs de factori primi în două moduri esenţial diferite şi pornind de la această ipoteză vom ajunge la o contradicţie. Această contradicţie va arăta că ipoteza existenţei unui întreg cu două descompuneri esenţial diferite în factori primi, este inacceptabilă şi, prin urmare, descompunerea în produs de factori primi, a oricărui întreg, este unică.

Dacă ar exista totuşi un întreg pozitiv, care ar putea fi descompus în două produse esenţial diferite de factori primi, va exista un cel mai mic întreg de acest fel (a se vedea p. 35).

(1) $$m = p_1 p_2 \ldots p_r = q_1 q_2 \ldots q_s,$$

unde p şi q reprezintă numere prime. Schimbînd ordinea acestor factori, dacă este necesar, putem presupune că

$$p_1 \leqslant p_2 \leqslant \ldots \leqslant p_r, \quad q_1 \leqslant q_2 \leqslant \ldots \leqslant q_s.$$

Observăm că p_1 nu poate fi egal cu q_1, deoarece, dacă ar fi, am putea şterge primul factor al fiecărui membru al egalităţii (1) şi am obţine două descompuneri esenţial diferite în factori primi, ale unui întreg mai mic decît m, ceea ce ar contrazice alegerea lui m, ca *cel mai mic* întreg pentru care acest lucru este posibil. Deci, sau $p_1 < q_1$ sau $q_1 < p_1$. Să presupunem că $p_1 < q_1$. (Dacă $q_1 < p_1$ vom schimba în cele ce urmează pe p şi q între ele.) Să formăm întregul

$$m' = m - (p_1 q_2 q_3 \ldots q_s). \tag{2}$$

Punînd în locul lui m cei doi membri ai egalităţii (1), putem scrie întregul m' sub una din următoarele două forme

$$m' = (p_1 p_2 \ldots p_r) - (p_1 q_2 \ldots q_s) = p_1(p_2 p_3 \ldots p_r - q_2 q_3 \ldots q_s) \tag{3}$$

$$m' = (q_1 q_2 \ldots q_s) - (p_1 q_2 \ldots q_s) = (q_1 - p_1)(q_2 q_3 \ldots q_s). \tag{4}$$

Deoarece $p_1 < q_1$, rezultă din (4) că m' este un întreg pozitiv, în timp ce din (2) rezultă că m' este mai mic decît m. Deci, descompunerea în factori primi a lui m' trebuie să fie *unică*, abstracţie făcînd de ordinea factorilor. Dar, din (3) rezultă că numărul prim p_1 este un factor al lui m', deci, din (4) reiese că p_1 trebuie să apară ca factor fie al lui $(q_1 - p_1)$, fie al lui $(q_2 q_3 \ldots q_s)$. (Aceasta rezultă din presupunerea unicităţii descompunerii în factori primi a numărului m'; a se vedea raţionamentul în paragraful următor.) Ultima alternativă este imposibilă, deoarece toate numerele q sînt mai mari decît p_1. Deci p_1 trebuie să fie un factor al lui $q_1 - p_1$, astfel încît, pentru un întreg h,

$$q_1 - p_1 = p_1 \cdot h \quad \text{sau} \quad q_1 = p_1(h + 1).$$

Dar aceasta arată că p_1 este un factor al lui q_1, ceea ce contrazice faptul că q_1 este număr prim. Contradicţia la care am ajuns arată că ipoteza noastră iniţială este inadmisibilă şi deci demonstraţia teoremei fundamentale a aritmeticii este completă.

Un corolar important al teoremei fundamentale este următorul: *Dacă un număr prim* p *este un factor al produsului* ab, *atunci* p *trebuie să fie factor fie al lui* a, *fie al lui* b. Într-adevăr, dacă p nu ar fi factor nici al lui a, nici al lui b, atunci produsul descompunerilor în factori primi, ale lui a şi b, ar da o descompunere în factori primi a întregului ab, *care nu conţine*

pe p. Pe de altă parte, dacă se presupune că *p* este un factor al lui *ab*, rezultă că există un întreg *t*, astfel încît

$$ab = pt.$$

Prin urmare, produsul lui *p* cu o descompunere în factori primi a lui *t* ar da o descompunere în factori primi a întregului *ab*, *care conține factorul p*, ceea ce contrazice unicitatea descompunerii în factori primi a lui *ab*.

Exemple: Dacă am stabilit că 13 este un factor al lui 2 652 și că 2 652 = = 6 · 442, putem conchide că 13 este un factor al lui 442. Pe de altă parte, 6 este un factor al lui 240, iar 240 = 15 · 16, dar 6 nu este factor nici al lui 15, nici al lui 16. De aici rezultă că ipoteza, că *p* este *număr prim* este esențială.

Exercițiu: Pentru a găsi toți divizorii unui număr *a* este suficient să-l descompunem pe *a* într-un produs de forma:

$$a = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_r^{\alpha_r},$$

unde factorii *p* sînt numere prime distincte, fiecare ridicat la o anumită putere. Toți divizorii lui *a* au forma

$$b = p_1^{\beta_1} \cdot p_2^{\beta_2} \cdots p_r^{\beta_r},$$

unde exponenții β sînt întregi oarecare, care satisfac inegalitățile

$$0 \leqslant \beta_1 \leqslant \alpha_1,\ 0 \leqslant \beta_2 \leqslant \alpha_2,\ \ldots,\ 0 \leqslant \beta_r \leqslant \alpha_r.$$

Demonstrați această afirmație. Ca o consecință, arătați că numărul divizorilor diferiți ai lui *a* (inclusiv divizorii *a* și 1) este egal cu

$$(\alpha_1 + 1)(\alpha_2 + 1) \ldots (\alpha_r + 1).$$

De exemplu,

$$144 = 2^4 \cdot 3^2$$

are 5 · 3 divizori. Ei sînt 1, 2, 4, 8, 16, 3, 6, 12, 24, 48, 9, 18, 36, 72, 144.

2. Distribuția numerelor prime

Se poate stabili o listă a tuturor numerelor prime pînă la un întreg dat *N* scriind în ordine crescătoare toți întregii mai mici decît *N*, ștergînd pe toți aceia care sînt multipli de 2, apoi pe toți aceia care au rămas și care sînt multipli de 3 și așa mai departe, pînă ce toate numerele compuse au fost eliminate. Acest procedeu, cunoscut sub numele „ciurul lui Eratostene”, prinde în ochiurile sale toate numerele prime de la 2 la *N*. Perfecționarea treptată a acestei metode a dus la stabilirea unor tabele complete de numere prime pînă la aproximativ 10 000 000. Ele ne furnizează o cantitate imensă de date empirice referitoare la distribuția și proprietățile numerelor prime.

Pe baza acestor tabele se pot face o serie de conjecturi foarte plauzibile — ca și cum teoria numerelor ar fi o știință experimentală. Adesea demonstrarea acestor ipoteze se arată a fi neobișnuit de grea.

a. *Formule care dau numere prime*

S-au făcut încercări de a găsi formule aritmetice simple, care să furnizeze numai numere prime, chiar dacă ele nu le dau pe toate. Fermat a făcut celebra conjectură (pe care nu a afirmat-o însă cu hotărîre) că toate numerele de forma

$$F(n) = 2^{2^n} + 1$$

sînt numere prime. Într-adevăr, pentru $n = 1, 2, 3, 4$, obținem

$$F(1) = 2^2 + 1 = 5$$
$$F(2) = 2^{2^2} + 1 = 2^4 + 1 = 17,$$
$$F(3) = 2^{2^3} + 1 = 2^8 + 1 = 257,$$
$$F(4) = 2^{2^4} + 1 = 2^{16} + 1 = 65\,537,$$

toate prime. Însă în 1732, Euler a descoperit factorizarea $2^{2^5} + 1 = 641 \cdot \cdot 6\,700\,417$; deci, numărul $F(5)$ nu este prim. Mai tîrziu s-a dovedit că și alte „numere Fermat" sînt numere compuse, fiind necesare metode mai profunde de teoria numerelor pentru fiecare caz, datorită dificultății insurmontabile a încercărilor directe. În momentul de față nu s-a demonstrat nici măcar faptul că vreunul din numerele $F(n)$ este prim pentru $n > 4$.

Iată o altă expresie simplă și remarcabilă care generează multe numere prime

$$f(n) = n^2 - n + 41.$$

Pentru $n = 1, 2, 3, \ldots, 40$, $f(n)$ este un număr prim; dar pentru $n = 41$ avem $f(n) = 41^2$, care nu mai este prim.

Expresia

$$n^2 - 79n + 1\,601$$

dă numere prime pînă la $n = 79$, dar nu și pentru $n = 80$. În general se poate spune că încercarea de a găsi formule elementare care să dea numai numere prime a fost zadarnică. Cu atît mai puțin promițătoare este încercarea de a găsi o formulă algebrică, care să dea *toate* numerele prime.

b. *Numerele prime în progresii aritmetice*

Dacă a fost simplu să se arate că există o infinitate de numere prime în șirul tuturor întregilor 1, 2, 3, 4, ..., saltul la șiruri, ca de pildă 1, 4, 7, 10, 13, ... sau 3, 7, 11, 15, 19, ... sau, mai general, la orice progresie

aritmetică a, $a + d$, $a + 2d$, ..., $a + nd$, ..., unde a şi d nu au nici ,su factor comun, a fost mult mai dificil. Toate observaţiile indicau numai faptul că *în orice astfel de progresie există o infinitate de numere prime*, ca şi în cea mai simplă dintre ele 1, 2, 3, ... A fost necesar un efort imens pentru demonstrarea acestei teoreme generale. Lejeune Dirichlet (1805—1859), unul dintre matematicienii de seamă ai secolului al XIX-lea, a obţinut un succe deplin aplicînd instrumentele cele mai avansate ale analizei matematice cunoscute pe vremea aceea. Lucrările sale originale, referitoare la acest subiect, se află chiar şi astăzi printre realizările remarcabile în matematică; după aproape o sută de ani, demonstraţia dată de el nu a fost încă suficient de simplificată pentru a fi înţeleasă şi de persoane care nu posedă în întregime analiza matematică şi teoria funcţiilor.

Nu vom încerca să demonstrăm aici teorema generală a lui Dirichlet şi ne vom limita la considerarea unei probleme ceva mai uşoare; vom generaliza demonstraţia lui Euclid că mulţimea numerelor prime este infinită, astfel încît să cuprindă cîteva progresii aritmetice speciale, ca de pildă $4n + 3$ sau $6n + 5$. Să ne oprim asupra celei dintîi. Observăm că orice număr prim mai mare decît 2 este impar (deoarece altfel ar fi divizibil cu 2) şi deci este de forma $4n + 1$ sau $4n + 3$, pentru un anumit întreg n. Mai departe, produsul a două numere de formă $4n + 1$ este din nou de aceeaşi formă, deoarece

$$(4a + 1)(4b + 1) = 16ab + 4a + 4b + 1 = 4(4ab + a + b) + 1.$$

Să presupunem că ar exista doar un număr finit de numere prime $p_1, p_2, \ldots, p_n$, de forma $4n + 3$ şi să considerăm numărul

$$N = 4(p_1p_2 \ldots p_n) - 1 = 4(p_1 \ldots p_n - 1) + 3.$$

Deci, sau N este un număr prim, sau el poate fi descompus într-un produs de numere prime, dintre care nici unul nu poate fi $p_1, \ldots, p_n$, deoarece aceste numere îl împart pe N cu restul -1. Mai mult, nu este posibil ca toţi factorii primi ai lui N să fie de forma $4n + 1$, deoarece N nu este de această formă şi, după cum am văzut, produsul unor numere de forma $4n + 1$ este de aceeaşi formă. Deci, cel puţin un factor prim trebuie să fie de forma $4n + 3$, ceea ce este imposibil, deoarece am văzut că nici unul dintre numerele p, pe care le-am presupus ca fiind *toate* numerele prime de forma $4n + 3$, nu poate fi un factor al lui N. Prin urmare, ipoteza că numărul numerelor prime de forma $4n + 3$ este finit duce la o contradicţie şi deci numărul numerelor prime de acest gen este infinit.

Exerciţiu: Demonstraţi teorema corespunzătoare pentru progresia $6n + 5$.

c. *Teorema numerelor prime*

În căutarea unei legi care guvernează distribuția numerelor prime, pasul decisiv a fost făcut în momentul în care matematicienii au renunțat la încercările zadarnice de a găsi o formulă matematică simplă, care să dea toate numerele prime, sau să dea numărul exact de numere prime conținute printre primii n întregi, și au căutat în schimb o informație referitoare la distribuția medie a numerelor prime printre întregi.

Pentru orice întreg n să notăm cu A_n numărul numerelor prime aflate printre întregii 1, 2, 3, ..., n. Dacă subliniem numerele prime aflate în șirul format de primii cîțiva întregi, 1, $\underline{2}$, $\underline{3}$, 4, $\underline{5}$, 6, $\underline{7}$, 8, 9, 10, $\underline{11}$, 12, $\underline{13}$, 14, 15, 16, $\underline{17}$, 18, $\underline{19}$, ..., putem calcula primele cîteva valori ale lui A_n:

$$A_1 = 0,\ A_2 = 1,\ A_3 = A_4 = 2,\ A_5 = A_6 = 3,\ A_7 = A_8 = A_9 = A_{10} = 4,$$

$$A_{11} = A_{12} = 5,\ A_{13} = A_{14} = A_{15} = A_{16} = 6,\ A_{17} = A_{18} = 7,\ A_{19} = 8 \text{ etc.}$$

Dacă luăm acum un șir oarecare de valori ale lui n, care crește nelimitat, de exemplu

$$n = 10,\ 10^2,\ 10^3,\ 10^4,\ \ldots$$

atunci valorile corespunzătoare ale lui

$$A_n\colon A_{10},\ A_{10^2},\ A_{10^3},\ A_{10^4},\ \ldots$$

vor crește, de asemenea, nelimitat (cu toate că mai încet). Întrucît știm că există o infinitate de numere prime, valorile lui A_n vor depăși, mai devreme sau mai tîrziu, orice număr finit. „Densitatea" numerelor prime printre primii n întregi este dată de raportul A_n/n; fără eforturi deosebite, cu ajutorul unor tabele de numere prime se pot calcula valorile lui A_n/n, pentru valori destul de mari ale lui n.

n	A_n/n
10^3	0,168
10^6	0,078498
10^9	0,050847478
....	....

Se poate considera că ultimul rînd al acestui tabel dă probabilitatea ca un întreg, ales la întîmplare din primii 10^9 întregi, să fie număr prim, deoarece în total există 10^9 alegeri posibile, dintre care A_{10^9} sînt numere prime.

Distribuția numerelor prime printre întregi este extrem de neregulată. Dar această neregularitate „în mic" dispare dacă ne fixăm atenția asupra distribuției medii a numerelor prime, ca fiind dată de raportul A_n/n. Legea simplă care guvernează comportarea acestui raport este una dintre cele mai remarcabile descoperiri din întreaga matematică. Pentru a enunța *teorema numerelor prime* trebuie să definim „logaritmul natural" al unui întreg n.

Pentru aceasta să luăm două axe perpendiculare într-un plan şi să considerăm locul geometric al tuturor punctelor din plan, pentru care produsul distanţelor x şi y la cele două axe este egal cu unu. Cu ajutorul coordonatelor x şi y, acest loc, o hiperbolă echilateră, este definită de ecuaţia $xy = 1$. Să definim acum log n ca fiind aria din fig. 5, mărginită de hiperbolă, axa

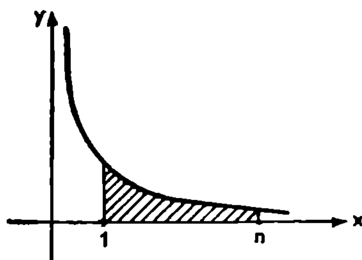


Fig. 5. Aria regiunii haşurate defineşte pe log n

ox, şi de cele două drepte verticale $x = 1$ şi $x = n$. (O discuţie mai detaliată referitoare la logaritm este dată în cap. VIII.) Din studiul empiric al tabelelor de numere prime, Gauss a observat că raportul A_n/n este aproximativ egal cu $1/\log n$, şi că această aproximaţie pare să se îmbunătăţească pe măsură ce n creşte. Ordinul de aproximaţie este dat de cîtul $\dfrac{A_n/n}{1/\log n}$, ale cărui valori pentru $n = 1\ 000,\ 1\ 000\ 000,\ 1\ 000\ 000\ 000$ sînt date în tabelul următor

n	A_n/n	$1/\log n$	$\dfrac{A_n/n}{1/\log n}$
10^3	0,168	0,145	1,159
10^6	0,078498	0,072382	1,084
10^9	0,050847478	0,048254942	1,053
. . .	. . .	. . .	. . .

Bazîndu-se pe aceste dovezi empirice, Gauss a emis conjectura că raportul A_n/n este „asimptotic egal" cu $1/\log n$. Prin aceasta se înţelege că dacă luăm un şir de valori din ce în ce mai mari ale lui n, de pildă

$$10,\ 10^2,\ 10^3,\ 10^4,\ \ldots$$

(cum am făcut şi mai înainte), atunci raportul dintre

$$\frac{A_n/n}{1/\log n},$$

calculat pentru aceste valori succesive ale lui n, se va apropia din ce în ce mai mult de 1, şi că diferenţa dintre acest cît şi 1 poate fi făcută oricît de mică vrem, limitîndu-ne la valori suficient de mari ale lui n. Această afirmaţie se exprimă simbolic, cu ajutorul semnului $\sim$:

$$\frac{A_n}{n} \sim \frac{1}{\log n} \text{ înseamnă că } \frac{A_n/n}{1/\log n} \text{ tinde spre 1 cînd } n \text{ creşte.}$$

Faptul că $\sim$ nu poate fi înlocuit cu semnul obişnuit al egalităţii ,,$=$'', este clar din faptul că, în timp ce A_n este întotdeauna un întreg, $n/\log n$ nu este întreg.

Faptul că comportarea medie a distribuţiei numerelor prime poate fi descrisă cu ajutorul funcţiei logaritmice, este o descoperire remarcabilă, deoarece este surprinzător ca două noţiuni matematice, care par atît de puţin înrudite, să fie în realitate atît de strîns legate.

Cu toate că afirmaţia din conjectura lui Gauss este uşor de înţeles, totuşi o demonstraţie matematică riguroasă a ei era peste puterile matematicii de pe vremea lui Gauss. Pentru a demonstra această teoremă, legată doar de cele mai elementare concepte, este necesar să folosim cele mai puternice metode ale matematicii moderne. Au trebuit aproape o sută de ani înainte ca analiza să se fi dezvoltat pînă la un punct la care Hadamard (1896), la Paris, şi de la Vallée Poussin (1896), la Louvain, au putut da o demonstraţie completă a teoremei numerelor prime. Simplificări şi modificări importante au fost date de Mangoldt şi Landau. Cu mult înainte de Hadamard, o muncă de pionierat decisivă a fost făcută de Riemann (1826—1866) într-o lucrare celebră, în care au fost trasate liniile strategice ale atacului. Matematicianul american Norbert Wiener a reuşit să modifice demonstraţia în aşa fel, încît să evite folosirea numerelor complexe într-un moment important al raţionamentului. Cu toate acestea, demonstraţia teoremei numerelor prime nu este de loc uşoară, chiar pentru un student avansat. Vom reveni asupra acestei probleme la p. 102 şi următoarele.

d. *Două probleme nerezolvate referitoare la numerele prime*

În timp ce problema distribuţiei medii a numerelor prime a fost rezolvată în mod satisfăcător, sînt multe alte conjecturi, care sînt confirmate de evidenţa empirică, dar care nu au fost încă demonstrate.

Una din acestea este celebra *conjectură a lui Goldbach.* Goldbach (1690—1764) nu a lăsat nici o urmă în istoria matematicii, cu excepţia acestei probleme, pe care el a propus-o, în 1742, într-o scrisoare către Euler. El a observat că în toate cazurile pe care le-a încercat, orice număr par (cu excepţia lui 2, care este el însuşi un număr prim) poate fi reprezentat ca suma a două numere prime. De exemplu, $4 = 2 + 2$, $6 = 3 + 3$, $8 = 5 + 3$, $10 = 5 + 5$, $12 = 5 + 7$, $14 = 7 + 7$, $16 = 13 + 3$, $18 = 11 + 7$, $20 = 13 + 7, \ldots$ $\ldots, 48 = 29 + 19, \ldots, 100 = 97 + 3$ etc.

Goldbach îl întreba pe Euler dacă ar putea demonstra că o astfel de reprezentare este posibilă pentru orice număr prim, sau dacă dimpotrivă ar putea găsi un exemplu, care să infirme această ipoteză. Euler nu a dat nici un răspuns şi nici nu l-a dat nimeni pînă în prezent. Evidenţa empirică în favoarea propoziţiei că orice număr par poate fi reprezentat în acest mod este pe deplin convingătoare, după cum poate verifica oricine, încercînd un număr de exemple. Sursa dificultăţii constă în faptul că numerele prime sînt definite cu ajutorul înmulţirii, în timp ce problema implică adunarea. În general, este foarte greu de stabilit legături între proprietăţile multiplicative şi cele aditive ale întregilor.

Pînă nu de mult, o demonstraţie a conjecturii lui Goldbach părea complet inabordabilă. Astăzi, soluţia nu mai pare de negăsit. Un succes important, foarte neaşteptat şi surprinzător pentru toţi experţii, a fost obţinut în 1931 de un tînăr matematician rus, necunoscut pe atunci, Schnirelmann (1905—1938), care a demonstrat că *orice întreg pozitiv poate fi reprezentat ca suma a cel mult* 300 000 *de numere prime*. Cu toate că acest rezultat pare ridicol în comparaţie cu scopul iniţial al demonstrării conjecturii lui Goldbach, el a constituit totuşi un prim pas în această direcţie. Demonstraţia este directă, constructivă, cu toate că ea nu furnizează nici o metodă practică pentru găsirea descompunerii unui întreg oarecare sub forma unei sume de numere prime. Mai tîrziu, matematicianul rus Vinogradov, folosind metode datorate lui Hardy, Littlewood şi marelui lor colaborator indian Ramanujan, a reuşit să reducă numărul de la 300 000 la 4. Aceasta este cu mult mai aproape de rezolvarea problemei lui Goldbach. Există însă o deosebire frapantă între rezultatul lui Schnirelmann şi cel al lui Vinogradov; mai semnificativă poate decît deosebirea dintre 300 000 şi 4. Teorema lui Vinogradov a fost demonstrată numai pentru toţi întregii „suficient de mari"; mai precis, Vinogradov a demonstrat că *există* un întreg N, astfel încît orice întreg $n > N$ poate fi reprezentat sub forma unei sume de cel mult patru numere prime. Metoda lui Vinogradov nu ne permite să apreciem pe N; spre deosebire de metoda lui Schnirelmann, ea este esenţial „indirectă" şi neconstructivă. Ceea ce a demonstrat într-adevăr Vinogradov este faptul că ipoteza existenţei unei infinităţi de numere întregi care nu pot fi descompuse sub forma unei sume de cel mult patru numere prime duce la contradicţii. Avem aici un exemplu concludent care arată deosebirea profundă dintre cele două tipuri de demonstraţie — directă şi indirectă (a se vedea discuţia generală de la p.102).

Următoarea problemă, mai surprinzătoare încă decît cea a lui Goldbach, nu s-a apropiat cîtuşi de puţin de rezolvare. S-a observat că numerele prime apar adesea în perechi de forma p şi $p + 2$. Astfel sînt 3 şi 5, 11 şi 13, 29 şi 31 etc. Se crede că afirmaţia existenţei unei infinităţi de astfel de perechi de numere este corectă, dar pînă în prezent nu s-a făcut nici cel mai mic pas spre a o demonstra.

§ 2. CONGRUENȚE

1. Noțiuni generale

Ori de cîte ori apare problema divizibilității întregilor printr-un întreg d fixat, noțiunea și notația de „congruență" (datorate lui Gauss) servesc la clarificarea și simplificarea raționamentului.

Pentru a introduce această noțiune, să examinăm resturile împărțirii întregilor prin numărul 5. Avem

$0 = 0 \cdot 5 + 0$	$7 = 1 \cdot 5 + 2$	$-1 = -1 \cdot 5 + 4$
$1 = 0 \cdot 5 + 1$	$8 = 1 \cdot 5 + 3$	$-2 = -1 \cdot 5 + 3$
$2 = 0 \cdot 5 + 2$	$9 = 1 \cdot 5 + 4$	$-3 = -1 \cdot 5 + 2$
$3 = 0 \cdot 5 + 3$	$10 = 2 \cdot 5 + 0$	$-4 = -1 \cdot 5 + 1$
$4 = 0 \cdot 5 + 4$	$11 = 2 \cdot 5 + 1$	$-5 = -1 \cdot 5 + 0$
$5 = 1 \cdot 5 + 0$	$12 = 2 \cdot 5 + 2$	$-6 = -2 \cdot 5 + 4$
$6 = 1 \cdot 5 + 1$	etc.	etc.

Observăm că restul împărțirii unui întreg prin 5 este unul din cele cinci numere 0, 1, 2, 3, 4. Spunem că doi întregi a și b sînt „congruenți modulo 5" dacă ei au același rest prin împărțirea cu 5. Astfel, 2, 7, 12, 17, 22, ..., -3, -8, -13, -18, ... sînt toate congruente modulo 5, deoarece ele dau restul 2. În general, spunem că doi întregi a și b sînt *congruenți modulo* d, unde d este un întreg fixat, dacă a și b dau același rest prin împărțirea cu d, cu alte cuvinte dacă există un întreg n, astfel încît $a - b = nd$. De exemplu, 27 și 15 sînt congruente modulo 4, deoarece

$$27 = 6 \cdot 4 + 3, \qquad 15 = 3 \cdot 4 + 3.$$

Noțiunea de congruență este atît de folositoare, încît e de dorit să avem pentru ea o notație cît mai scurtă. Scriem

$$a \equiv b \pmod{d}$$

pentru a exprima faptul că a și b sînt congruente modulo d. Dacă nu există nici o îndoială referitoare la modul, atunci „mod d" din formulă poate fi omis. (Dacă a nu este congruent cu b modulo d, vom scrie $a \not\equiv b \pmod{d}$.)

Congruențele apar frecvent în viața de toate zilele. De exemplu, acele unui ceasornic indică ora modulo 12, iar indicatorul de kilometraj al unui automobil dă numărul total de kilometri parcurși, modul 100 000.

Înainte de a trece la o expunere detaliată a congruențelor și proprietăților lor, cititorul ar trebui să verifice că următoarele propoziții sînt într-adevăr echivalente.

1. a este congruent cu b modulo d.
2. $a = b + nd$, pentru un întreg n, convenabil.
3. d divide pe $a - b$.

Introducerea notației lui Gauss pentru congruență subliniază faptul că congruența are multe din proprietățile formale ale egalității obișnuite. Cele mai importante dintre ele sînt următoarele:

1) Întotdeauna $a = a$.
2) Dacă $a = b$, atunci $b = a$.
3) Dacă $a = b$ și $b = c$, atunci $a = c$.

Mai mult, dacă $a = a'$ și $b = b'$, atunci

4) $a + b = a' + b'$.
5) $a - b = a' - b'$.
6) $ab = a'b'$.

Aceste proprietăți rămîn adevărate dacă relația $a = b$ este înlocuită prin relația de congruență $a \equiv b \pmod{d}$. Astfel,

1') Întotdeauna $a \equiv a \pmod{d}$.
2') Dacă $a \equiv b \pmod{d}$, atunci $b \equiv a \pmod{d}$.
3') Dacă $a \equiv b \pmod{d}$ și $b \equiv c \pmod{d}$, atunci $a \equiv c \pmod{d}$.

Verificarea acestor proprietăți este lăsată pe seama cititorului.
Mai mult, dacă $a \equiv a' \pmod{d}$ și $b \equiv b' \pmod{d}$, atunci

4') $a + b \equiv a' + b' \pmod{d}$
5') $a - b \equiv a' - b' \pmod{d}$
6') $ab \equiv a'b' \pmod{d}$.

Astfel, *congruențele în raport cu același modul pot fi adunate, scăzute și înmulțite.* Pentru a demonstra aceste trei afirmații, trebuie să observăm numai că dacă

$$a = a' + rd, \quad b = b' + sd,$$

atunci

$$a + b = a' + b' + (r + s)d,$$
$$a - b = a' - b' + (r - s)d,$$
$$ab = a'b' + (a's + b'r + rsd)d,$$

din care rezultă concluziile dorite.

Noțiunea de congruență are o minunată interpretare geometrică. De obicei, dacă dorim să dăm o reprezentare geometrică numerelor întregi, alegem un segment de dreaptă drept unitate de lungime și apoi îl extindem prin multiplii săi, în ambele sensuri. În acest mod putem găsi pentru orice număr întreg un punct corespunzător pe dreaptă (fig. 6). Dar, dacă vrem să tra-

tăm întregii modulo d, orice două numere congruente sînt considerate identice, în ceea ce priveşte comportarea lor prin împărţirea cu a, deoarece ele au acelaşi rest. Pentru a reprezenta geometric acest lucru, luăm un cerc pe care îl împărţim în d părţi egale. Orice întreg împărţit prin d are ca rest unul din cele d numere 0, 1, ..., $d - 1$, care sînt aşezate la distanţe



Fig. 6. Reprezentarea geometrică a întregilor

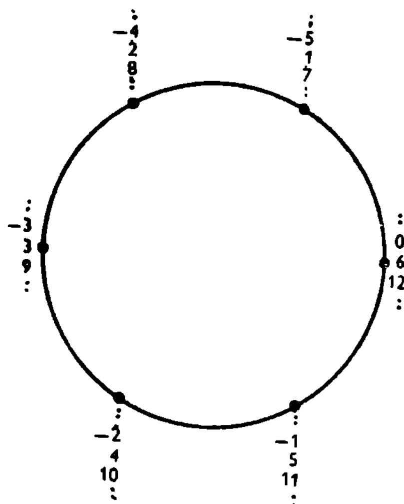


Fig. 7. Reprezentarea geometrică a întregilor modulo 6

egale pe circumferinţa cercului. Orice întreg este congruent modulo d cu unul din aceste numere şi deci este reprezentat geometric printr-unul din aceste puncte; două numere sînt congruente dacă ele sînt reprezentate de acelaşi punct. Fig. 7 ilustrează cazul $d = 6$. Cadranul unui ceasornic constituie un alt exemplu din viaţa de toate zilele.

Ca exemplu de folosire a proprietăţii multiplicative 6') a congruenţelor, putem determina resturile împărţirii succesive a puterilor lui 10, printr-un număr dat. De exemplu,

$$10 \equiv -1 \pmod{11},$$

deoarece $10 = -1 + 11$. Înmulțind succesiv această congruență prin ea însăși obținem

$$10^2 \equiv (-1)(-1) = 1 \pmod{11},$$
$$10^3 \equiv -1 \quad ,, \quad ,$$
$$10^4 \equiv 1 \quad ,, \quad \text{etc.}$$

De aici putem arăta că orice întreg

$$z = a_0 + a_1 \cdot 10 + a_2 \cdot 10^2 + \ldots + a_n \cdot 10^n,$$

exprimat în sistemul zecimal, are același rest prin împărțirea cu 11 ca și suma cifrelor sale, luată cu semne alternate,

$$t = a_0 - a_1 + a_2 - a_3 + \ldots.$$

Într-adevăr, avem

$$z - t = a_1 \cdot 11 + a_2(10^2 - 1) + a_3(10^3 + 1) + a_4(10^4 - 1) + \ldots.$$

Deoarece toate numerele $11, 10^2 - 1, 10^3 + 1, \ldots$ sînt congruente cu 0 modulo 11, $z - t$ are aceeași proprietate, și de aceea z are același rest prin împărțirea cu 11 ca și t. Rezultă, în particular, că un număr se divide cu 11 (are restul 0) dacă și numai dacă suma alternată a cifrelor sale este divizibilă cu 11. De exemplu, numărul $z = 3\,162\,819$ se divide cu 11, deoarece $3 - 1 + 6 - 2 + 8 - 1 + 9 = 22$. Găsirea unei reguli de divizibilitate cu 3 sau 9 este un lucru mai simplu, deoarece $10 \equiv 1$ (mod 3 sau 9) și de aceea $10^n \equiv 1$ (mod 3 sau 9) pentru orice n. Rezultă că un număr z se divide cu 3 sau 9 dacă și numai dacă suma cifrelor sale

$$s = a_0 + a_1 + a_2 + \ldots + a_n$$

se divide respectiv cu 3 sau 9.

Pentru congruențele modulo 7 avem

$$10 \equiv 3,\ 10^2 \equiv 2,\ 10^3 \equiv -1,\ 10^4 \equiv -3,\ 10^5 \equiv -2,\ 10^6 \equiv 1.$$

Mai departe, resturile se repetă. Astfel, z se divide cu 7 dacă și numai dacă expresia

$$r = a_0 + 3a_1 + 2a_2 - a_3 - 3a_4 - 2a_5 + a_6 + 3a_7 + \ldots$$

este divizibilă cu 7.

Exercițiu: Găsiți o regulă similară pentru divizibilitatea cu 13.

Adunînd sau înmulțind congruențele față de un modul dat, de exemplu, $d = 5$, putem evita ca numerele care intervin să devină prea mari, înlocuind întotdeauna orice număr a prin numărul din mulțimea

$$0,\ 1,\ 2,\ 3,\ 4$$

cu care el este congruent. Astfel, pentru a calcula sume și produse de întregi modulo 5, trebuie să folosim doar următoarele tabele de adunare și înmulțire.

$a + b$

$b \equiv$	0	1	2	3	4
$a \equiv 0$	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$a \cdot b$

$b \equiv$	0	1	2	3	4
$a \equiv 0$	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Din al doilea tabel se vede că un produs ab este congruent cu 0 (mod 5) numai dacă a sau $b \equiv 0$ (mod 5). Acest lucru sugerează legea generală

7) $ab \equiv 0 \pmod{d}$ numai dacă $a \equiv 0$ sau $b \equiv 0 \pmod{d}$,

care este o extindere a legii obișnuite a întregilor, care afirmă că $ab = 0$ numai dacă $a = 0$ sau $b = 0$. *Legea* 7) *este valabilă numai dacă modulul* d *este un număr prim.* Într-adevăr, congruența

$$ab \equiv 0 \pmod{d}$$

înseamnă că d divide pe ab, și am văzut că un număr prim d divide un produs ab, numai dacă el divide pe a sau pe b; adică numai dacă

$$r \equiv 0 \pmod{d} \quad \text{sau} \quad b \equiv 0 \pmod{d}.$$

Dacă d nu este număr prim, legea nu mai este valabilă; pentru că putem scrie $d = r \cdot s$, unde r și s sînt mai mici decît d, astfel încît

$$r \not\equiv 0 \pmod{d}, \quad s \not\equiv 0 \pmod{d},$$

însă

$$rs = d \equiv 0 \pmod{d}.$$

De exemplu $2 \not\equiv 0 \pmod{6}$ și $3 \not\equiv 0 \pmod{6}$, însă $2 \cdot 3 = 6 \equiv 0 \pmod{6}$.

Exercițiu: Arătați că este valabilă următoarea *lege de simplificare* pentru congruențele în raport cu un modul prim:

Dacă

$$ab \equiv ac \text{ și } a \not\equiv 0, \text{ atunci } b \equiv c.$$

Exerciții: 1) Cu care număr, cuprins între 0 și 6 inclusiv, este congruent modulo 7, produsul $11 \cdot 18 \cdot 2322 \cdot 13 \cdot 19$?

2) Cu ce număr, cuprins între 0 și 12 inclusiv, este congruent modulo 13 numărul $3 \cdot 7 \cdot 11 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 113$?

3) Cu ce număr, cuprins între 0 și 4 inclusiv, este congruentă modulo 5 suma $1 + 2 + 2^2 + \ldots + 2^{19}$?

2. Teorema lui Fermat

În secolul al XVII-lea, Fermat, fondatorul teoriei moderne a numerelor, a descoperit o teoremă foarte importantă: *Dacă* p *este un număr prim care nu divide întregul* a, *atunci*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Aceasta înseamnă că puterea $(p - 1)$-a a numărului a dă restul 1 prin împărțirea cu p.

Cîteva din calculele noastre precedente confirmă această teoremă; de exemplu, am găsit că $10^6 \equiv 1 \pmod 7$, $10^2 \equiv 1 \pmod 3$, și $10^{10} \equiv 1 \pmod{11}$. În mod similar putem arăta că $2^{12} \equiv 1 \pmod{13}$ și $5^{10} \equiv 1 \pmod{11}$. Pentru a verifica ultimele congruențe, nu este necesar să calculăm într-adevăr puteri atît de mari, deoarece putem folosi proprietatea multiplicativă a congruențelor

$$\begin{aligned} 2^4 &= 16 \equiv 3 \pmod{13}, \\ 2^8 &= 9 \equiv -4 \quad ,, \quad , \\ 2^{12} &= -4 \cdot 3 = -12 \equiv 1 \pmod{13}, \end{aligned}$$

$$\begin{aligned} 5^2 &\equiv 3 \pmod{11}, \\ 5^4 &\equiv 9 \equiv -2 \quad ,, \quad , \\ 5^8 &\equiv 4 \pmod{11}, \\ 5^{10} &\equiv 3 \cdot 4 = 12 \equiv 1 \pmod{11}. \end{aligned}$$

Pentru a demonstra teorema lui Fermat, considerăm multiplii lui a

$$m_1 = a,\ m_2 = 2a,\ m_3 = 3a,\ \ldots,\ m_{p-1} = (p - 1)a.$$

Oricare doi din acești întregi nu pot fi congruenți modulo p între ei, deoarece atunci p ar fi un factor al lui $m_r - m_s = (r - s)a$, pentru o anumită pereche de întregi r, s cu $1 \leqslant r < s \leqslant (p - 1)$. Însă legea 7) arată că acest lucru nu se poate întîmpla, deoarece dacă $s - r$ este mai mic decît p, p nu este un factor al lui $s - r$, în timp ce, prin ipoteză, p nu este un factor al lui a. În mod similar, nici unul din aceste numere nu poate fi congruent cu 0. De aici rezultă că numerele $m_1, m_2, \ldots, m_{p-1}$ trebuie să fie respectiv congruente cu numerele $1, 2, 3, \ldots, p - 1$, într-o anumită ordine. Conchidem că

$$m_1 m_2 \ldots m_{p-1} = 1 \quad 2 \cdot 3 \ldots (p-1)\, a^{p-1} \equiv 1 \cdot 2 \cdot 3 \ldots (p-1) \pmod{p},$$

sau, dacă scriem pentru prescurtare $K = 1 \cdot 2 \cdot 3 \ldots (p - 1)$,

$$K(a^{p-1} - 1) \equiv 0 \pmod{p}.$$

Dar K nu se divide prin p, deoarece nici unul din factorii lui nu se divide prin p; deci, în virtutea legii 7), $(a^{p-1} - 1)$ trebuie să fie divizibil cu p, adică

$$a^{p-1} - 1 \equiv 0 \pmod{p}.$$

Aceasta este teorema lui Fermat.

Pentru a verifica această teoremă încă o dată, să luăm $p = 23$ și $a = 5$. Atunci avem, modulo 23, $5^2 \equiv 2$, $5^4 \equiv 4$, $5^8 \equiv 16 \equiv -7$, $5^{16} \equiv 49 \equiv 3$, $5^{20} \equiv 12$, $5^{22} \equiv 24 \equiv 1$. Dacă luăm $a = 4$ în locul lui 5, găsim, tot modulo 23, $4^2 \equiv -7$, $4^3 \equiv -28 \equiv -5$, $4^4 \equiv -20 \equiv 3$, $4^6 \equiv 9$, $4^{11} \equiv -45 \equiv 1$, $4^{22} \equiv 1$.

În exemplul de mai sus cu $a = 4$, $p = 23$, ca și în altele, observăm că nu numai puterea $(p-1)$—a a lui a, dar și o putere mai mică, poate fi congruentă cu 1. Cea mai mică putere de acest fel, în cazul nostru 11, este întotdeauna un divizor al lui $p - 1$. (A se vedea exercițiul 3 de mai jos).

Exerciții: 1) Arătați printr-un calcul similar că $2^8 \equiv 1 \pmod{17}$; $3^8 \equiv -1 \pmod{17}$; $3^{14} \equiv -1 \pmod{29}$; $2^{14} \equiv -1 \pmod{29}$; $4^{14} \equiv 1 \pmod{29}$; $5^{14} \equiv 1 \pmod{29}$.

2) Verificați teorema lui Fermat pentru $p = 5, 7, 11, 17$ și 23, dînd diferite valori lui a.

3) Demonstrați teorema generală: Cel mai mic întreg pozitiv e pentru care $a^e \equiv 1 \pmod{p}$ trebuie să fie un divizor al lui $p - 1$. (Indicație: Împărțiți pe $p - 1$ prin e, obțineți

$$p - 1 = ke + r,$$

unde $0 \leqslant r < e$, și mai departe folosiți faptul că $a^{p-1} \equiv a^e \equiv 1$ (mod p).)

3. Resturi pătratice

Întorcîndu-ne din nou la exemplele care ilustrau teorema lui Fermat, observăm nu numai că avem întotdeauna $a^{p-1} \equiv 1 \pmod{p}$, dar (dacă p este număr prim, diferit de 2 și deci impar, de forma $p = 2p' + 1$) pentru anumite valori ale lui a, $a^{p'} = a^{(p-1)/2} \equiv 1 \pmod{p}$. Acest fapt sugerează un lanț de investigații interesante. Putem scrie teorema lui Fermat sub forma următoare:

$$a^{p-1} - 1 = a^{2p'} - 1 = (a^{p'} - 1)(a^{p'} + 1) \equiv 0 \pmod{p}.$$

Deoarece un produs se divide cu p numai dacă unul din factorii lui se divide cu p, rezultă imediat că sau $a^{p'} - 1$ sau $a^{p'} + 1$ trebuie să fie divizibil cu p, astfel încît pentru orice număr prim $p > 2$ și orice număr a, nedivizibil cu p, avem sau

$$a^{(p-1)/2} \equiv 1 \text{ sau } a^{(p-1)/2} \equiv -1 \pmod{p}.$$

Încă de la începutul teoriei moderne a numerelor, matematicienii s-au interesat de găsirea numerelor a pentru care are loc prima congruență, cît și a numerelor pentru care are loc cea de-a doua. Să presupunem că a este congruent modulo p cu pătratul unui număr x,

$$a \equiv x^2 \pmod{p}.$$

Atunci $a^{(p-1)/2} \equiv x^{p-1}$, conform teoremei lui Fermat, membrul drept, prin urmare și cel din stînga, este congruent cu 1 $\pmod{p}$. Un număr a, nedivizibil cu p, care este congruent modulo p cu pătratul unui număr, se numește

rest pătratic al lui p, în timp ce un număr b, nemultiplu de p, care nu este congruent cu nici un pătrat, se numește *non-rest pătratic* al lui p. Am văzut puțin mai înainte că orice rest pătratic a al lui p satisface congruența $a^{(p-1)/2} \equiv 1 \pmod p$. Se poate arăta cu ușurință că pentru orice non-rest b avem congruența $b^{(p-1)/2} \equiv -1 \pmod p$. Mai mult, vom arăta acum că printre numerele $1, 2, 3, \ldots, p-1$ se află exact $(p-1)/2$ resturi și $(p-1)/2$ non-resturi pătratice.

Cu toate că multe date empirice pot fi obținute prin calcul direct, la început nu a fost ușor de descoperit legi generale care guvernează distribuția resturilor și non-resturilor pătratice. Prima proprietate fundamentală a acestor resturi a fost observată de Legendre (1752—1833), și numită mai tîrziu de Gauss *legea reciprocității pătratice*. Această lege se referă la comportarea a două numere prime diferite p și q, și afirmă că q este rest pătratic al lui p, dacă și numai dacă p este rest pătratic al lui q, cu condiția ca produsul $\left(\frac{p-1}{2}\right) \cdot \left(\frac{q-1}{2}\right)$ să fie *par*. În cazul în care acest produs este *impar*, situația se răstoarnă, p este un rest al lui q, dacă și numai dacă q este *non-rest* al lui p. Una din realizările tînărului Gauss a fost prima demonstrație riguroasă a acestei teoreme remarcabile, care multă vreme a provocat pe matematicieni. Demonstrația dată de Gauss nu a fost cîtuși de puțin simplă, iar legea reciprocității nu este ușor de stabilit nici astăzi, cu toate că au fost publicate numeroase alte demonstrații. Adevărata ei semnificație a apărut în legătură cu dezvoltările moderne din teoria algebrică a numerelor.

Ca exemplu ilustrativ al distribuției resturilor pătratice, să alegem $p = 7$. Atunci, deoarece

$$0^2 \equiv 0,\ 1^2 \equiv 1,\ 2^2 \equiv 4,\ 3^2 \equiv 2,\ 4^2 \equiv 2,\ 5^2 \equiv 4,\ \ 6^2 \equiv 1,$$

toate modulo 7, și deoarece pătratele următoare repetă acest șir, resturile pătratice ale lui 7 sînt numerele congruente cu 1, 2 sau 4, în timp ce non-resturile sînt congruente cu 3, 5 sau 6. În cazul general, resturile pătratice ale lui p sînt numerele congruente cu $1^2, 2^2, \ldots, (p-1)^2$. Dar acestea sînt congruente cîte două, deoarece

$$x^2 \equiv (p-x)^2 \pmod p \quad (\text{de exemplu } 2^2 \equiv 5^2 \pmod 7),$$

pentru că $(p-x)^2 = p^2 - 2px + x^2 \equiv x^2 \pmod p$. Deci jumătate din numerele $1, 2, \ldots, p-1$ sînt resturi pătratice ale lui p, iar jumătate sînt non-resturi.

Pentru a ilustra legea reciprocității pătratice, să alegem $p = 5$, $q = 11$. Deoarece $11 \equiv 1^2 \pmod 5$, 11 este un rest pătratic (mod 5); deoarece produsul $[(5-1)/2][(11-1)/2]$ este par, legea reciprocității ne spune că 5 este un rest pătratic (mod 11). Ca o confirmare a acestui fapt observăm că

$5 \equiv 4^2 \pmod{11}$. Pe de altă parte, dacă $p = 7$, $q = 11$, produsul $[(7-1)/2] \cdot [(11 - 1/2)]$ este impar și, într-adevăr, 11 este un rest (mod 7) (deoarece $11 \equiv 2^2 \pmod 7$), în timp ce 7 este non-rest (mod 11).

Exerciții: 1. $6^2 = 36 \equiv 13 \pmod{23}$. Este 23 rest pătratic (mod 13)?

2. Am văzut că $x^2 \equiv (p - x)^2 \pmod p$. Arătați că acestea sînt *singurele* congruențe printre numerele $1^2, 2^2, 3^2, \ldots, (p-1)^2$.

§ 3. NUMERELE PITAGOREICE ȘI ULTIMA TEOREMĂ A LUI FERMAT

O problemă interesantă din teoria numerelor este legată de teorema lui Pitagora. Grecii știau că un triunghi cu laturile 3, 4, 5 este dreptunghic. Aceasta sugerează întrebarea generală: Care alte triunghiuri dreptunghice au laturi, ale căror lungimi sînt multipli întregi ai unității de lungime? Teorema lui Pitagora se exprimă algebric prin ecuația

$$a^2 + b^2 = c^2, \tag{1}$$

unde a și b sînt lungimile catetelor, iar c este lungimea ipotenuzei. Problema găsirii *tuturor* triunghiurilor dreptunghice, ale căror laturi au lungimi întregi, este astfel echivalentă cu problema găsirii tuturor soluțiilor întregi (a, b, c) ale ecuației (1). Orice astfel de triplet de numere se numește *triplet numeric pitagoreic*.

Problema găsirii tuturor tripletelor numerice pitagoreice poate fi rezolvată foarte simplu. Dacă a, b și c formează un triplet numeric pitagoreic, astfel încît $a^2 + b^2 = c^2$, atunci să notăm pentru prescurtare, $a/c = x$, $b/c = y$, unde x și y sînt numere raționale, pentru care $x^2 + y^2 = 1$. Atunci avem $y^2 = (1 - x)(1 + x)$, sau $y/(1 + x) = (1 - x)/y$. Valoarea comună a celor doi membri ai acestei ecuații este un număr t, care se poate exprima sub forma unui cît de doi întregi, u/v. Putem scrie acum $y = t(1 + x)$ și $(1 - x) = ty$, sau

$$tx - y = -t, \quad x + ty = 1.$$

Din acest sistem de ecuații deducem imediat că

$$x = \frac{1 - t^2}{1 + t^2}, \quad y = \frac{2t}{1 + t^2}.$$

Făcînd substituțiile obținem

$$\frac{a}{c} = \frac{v^2 - u^2}{u^2 + v^2}, \quad \frac{b}{c} = \frac{2uv}{u^2 + v^2}.$$

De aceea,

$$a = (v^2 - u^2)r, \quad b = (2uv)r, \quad c = (u^2 + v^2)r, \tag{2}$$

unde r este un factor de proporționalitate rațional. Aceasta arată că dacă (a, b, c) este un triplet numeric pitagoreic, atunci a, b, c sînt proporționale respectiv cu $v^2 - u^2$, $2uv$, $u^2 + v^2$. Reciproc, este ușor de văzut că orice triplet (a, b, c) definit de (2) este un triplet pitagoreic, deoarece din (2) obținem

$$a^2 = (u^4 - 2u^2v^2 + v^4)r^2,$$
$$b^2 = (4u^2v^2)r^2,$$
$$c^2 = (u^4 + 2u^2v^2 + v^4)r^2,$$

astfel încît $a^2 + b^2 = c^2$.

Acest rezultat mai poate fi încă simplificat. Din orice triplet numeric pitagoreic (a, b, c) putem deduce un număr infinit de alte triplete pitagoreice (sa, sb, sc), pentru orice întreg pozitiv s. Astfel, din (3, 4, 5) putem obține (6, 8, 10), (9, 12, 15) etc. Prin urmare, asemenea triplete nu sînt esențial distincte, deoarece ele corespund unor triunghiuri dreptunghice asemenea. De aceea, vom spune că un triplet numeric pitagoreic este *primitiv* dacă numerele a, b, c nu au nici un factor comun. Se poate arăta că *formulele*

$$a = v^2 - u^2,$$
$$b = 2uv,$$
$$c = u^2 + v^2,$$

pentru orice întregi pozitivi u *și* v, *cu* v > u, *unde* u *și* v *nu au nici un factor comun și nu sînt ambele impare, dau toate tripletele numerice pitagoreice primitive.*

**Exercițiu*: Demonstrați ultima afirmație.

Iată cîteva exemple de triplete numerice pitagoreice primitive:

$$u = 2, v = 1 : (3, 4, 5), \; u = 3, v = 2 : (5, 12, 13), \; u = 4, v = 3 : (7, 24, 25), \ldots, u = 10, v = 7 : (51, 140, 149) \text{ etc.}$$

Acest rezultat referitor la numerele pitagoreice pune în mod natural problema găsirii unor întregi a, b, c pentru care să avem $a^3 + b^3 = c^3$, sau $a^4 + b^4 = c^4$, sau, în general, dacă pentru un exponent întreg pozitiv dat $n > 2$, ecuația

$$a^n + b^n = c^n \tag{3}$$

poate fi rezolvată cu întregi pozitivi a, b, c. Răspunsul spectaculos a fost dat de Fermat. Studiind opera lui Diofant, matematician din antichitate, care a avut contribuții remarcabile în teoria numerelor, Fermat și-a făcut însemnări pe marginea paginilor exemplarului său. Cu toate că a enunțat în acest fel numeroase teoreme, fără a le demonstra însă, ulterior s-au găsit demonstrații ale tuturor teoremelor cuprinse în notele marginale, cu o singură excepție importantă. Comentînd numerele pitagoreice, Fermat a afirmat că *ecuația* (3) *nu este rezolvabilă în numere întregi, oricare ar fi* $n > 2$, dar că, din nefericire, demonstrația elegantă pe care a găsit-o era prea lungă pentru ca să încapă pe marginea foii pe care lucra.

Afirmația generală a lui Fermat nu a fost dovedită sau infirmată niciodată, în ciuda eforturilor unui șir de matematicieni renumiți. Teorema a fost într-adevăr demonstrată, pentru numeroase valori ale lui n, în particular pentru toți $n < 619$, însă nu pentru orice n, cu toate că nu a fost găsit nici un contraexemplu. Cu toate că teorema, în sine, nu este atît de importantă din punct de vedere matematic, încercările de a o demonstra au dat naștere unor investigații importante în teoria numerelor. Problema a suscitat mult interes și în cercuri mai largi — în parte datorită unui premiu de 100 000 de mărci, oferit persoanei care va da prima soluție, și păstrat de Academia Regală din Göttingen. Pînă ce inflația postbelică din Germania nu a dus la devalorizarea monetară a acestui premiu, numeroase „soluții" incorecte erau prezentate în fiecare an Academiei. Chiar matematicieni serioși se înșelau uneori, înmînînd sau publicînd demonstrații care cădeau după descoperirea unei greșeli superficiale. Interesul general pentru această problemă pare să fi scăzut după devalorizarea mărcii, cu toate că, din timp în timp, se mai anunță în presă că problema a fost rezolvată de „un geniu", necunoscut pînă atunci.

§ 4. ALGORITMUL LUI EUCLID

1. Teoria generală

Cititorul este familiarizat cu algoritmul obișnuit al împărțirii unui întreg a prin alt întreg b și știe că acest procedeu poate fi continuat pînă ce restul devine mai mic decît împărțitorul. Astfel, dacă $a = 648$ și $b = 7$, avem un cît $q = 92$ și un rest $r = 4$.

```
648 | 7
63  |----
--- | 92
 18
 14
 --
  4
```

$$648 = 7 \cdot 92 + 4.$$

Putem enunța aceasta sub forma unei teoreme generale: *Dacă* a *este un întreg oarecare și* b *este un întreg mai mare decît* 0, *atunci putem găsi întotdeauna un întreg* q, *astfel încît*

(1) $$a = b \cdot q + r,$$

unde r *este un întreg, care satisface inegalitatea* $0 \leqslant r < b$.

Demonstrăm această propoziție, fără a folosi algoritmul împărțirii. Este suficient să observăm că orice întreg a este fie el însuși un multiplu de b,

$$a = bq,$$

fie se află între doi multipli succesivi ai lui b,

$$bq < a < b\ (q + 1) = bq + b.$$

În primul caz, egalitatea (1) are loc cu $r = 0$. În al doilea caz, avem, din prima din inegalitățile de mai sus

$$a - bq = r > 0,$$

iar din a doua

$$a - bq = r < b,$$

astfel încît $0 < r < b$, așa cum se cere în (1).

Din acest rezultat simplu vom deduce numeroase consecințe importante. Prima din acestea este o metodă pentru găsirea celui mai mare divizor comun a două numere întregi.

Fie a și b doi întregi oarecare, nu ambii nuli; să considerăm mulțimea tuturor întregilor pozitivi, care divid atît pe a, cît și pe b. Această mulțime este cu siguranță finită, deoarece dacă, de exemplu, $a \neq 0$, atunci nici un întreg mai mare decît a nu poate fi divizor al lui a, pentru a nu mai vorbi de b. Prin urmare, există numai un număr finit de divizori comuni ai lui a și b; să notăm cu d pe cel mai mare dintre ei. Întregul d se numește *cel mai mare divizor comun* al lui a și b, și se notează $d = (a,b)$. Astfel, pentru $a = 8$ și $b = 12$, găsim prin încercare directă că $(8, 12) = 4$, în timp ce pentru $a = 5$ și $b = 9$ găsim că $(5, 9) = 1$. Dacă a și b sînt suficient de mari, de exemplu, $a = 1\,804$ și $b = 328$, încercarea de a găsi pe (a, b) este destul de laborioasă. O metodă scurtă și sigură este dată de *algoritmul lui Euclid.* (Un algoritm este o metodă sistematică de calcul.) Algoritmul lui Euclid se bazează pe faptul că din orice relație de forma

(2) $$a = b \cdot q + r$$

rezultă că

(3) $$(a, b) = (b, r).$$

Într-adevăr, orice număr u, care divide atît pe a cît și pe b,

$$a = su, \qquad b = tu,$$

divide de asemenea și pe r, deoarece $r = a - bq = su - qtu = (s - qt)u$; și reciproc, orice număr v, care divide pe b și pe r,

$$b = s'v, \qquad r = t'v,$$

divide de asemenea și pe a, deoarece $a = bq + r = s'vq + t'v = (s'q + t')v$. Prin urmare, *orice* divizor comun al lui a și b este, în același timp, un divizor comun al lui b și r, și reciproc. Deoarece mulțimea *tuturor* divizorilor comuni ai lui a și b este identică cu mulțimea tuturor divizorilor comuni ai lui b și r, *cel mai mare divizor comun* al lui a și b trebuie să fie egal cu cel mai mare divizor comun al lui b și r, ceea ce stabilește egalitatea (3). Utilitatea acestei relații se va vedea imediat.

Să revenim la problema găsirii celui mai mare divizor comun al lui 1 804 și 328. Prin împărțirea obișnuită

$$\begin{array}{r|l} 1804 & 328 \\ \cline{2-2} 1640 & 5 \\ \cline{1-1} 164 & \end{array}$$

găsim că

$$1804 = 5 \cdot 328 + 164.$$

Deci, din (3) găsim că

$$(1804,\ 328) = (328,\ 164).$$

Observăm că problema găsirii celui mai mare divizor comun al lui 1804 și 328 a fost înlocuită printr-o problemă analogă, care depinde însă de numere mai mici. Putem continua procedeul. Deoarece

$$\begin{array}{r|l} 328 & 164 \\ \cline{2-2} 328 & 2 \\ \cline{1-1} 0 & \end{array},$$

avem $328 = 2 \cdot 164 + 0$, astfel încît $(328,\ 164) = (164,\ 0) = 164$. Deci $(1804,\ 328) = (328,\ 164) = (164,\ 0) = 164$, care este rezultatul căutat.

Acest procedeu de găsire a celui mai mare divizor comun a două numere este dat într-o formă geometrică în *Elementele* lui Euclid. Pentru întregi arbitrari a și b, nu ambii nuli, el poate fi descris aritmetic în modul următor.

Putem presupune că $b \neq 0$, deoarece $(a, 0) = a$. Atunci, prin împărţiri succesive putem scrie

$$(4) \qquad \begin{aligned} a &= bq_1 + r_1 \qquad (0 < r_1 < b) \\ b &= r_1q_2 + r_2 \qquad (0 < r_2 < r_1) \\ r_1 &= r_2q_3 + r_3 \qquad (0 < r_3 < r_2) \\ r_2 &= r_3q_4 + r_4 \qquad (0 < r_4 < r_3) \\ &\dots\dots\dots\dots\dots\dots \end{aligned}$$

atîta timp cît resturile $r_1, r_2, r_3, \dots$, nu sînt nule. Privind inegalităţile care se află în dreapta, vedem că resturile succesive formează un şir strict descrescător de numere pozitive:

$$(5) \qquad b > r_1 > r_2 > r_3 > r_4 > \dots > 0.$$

Deci, după cel mult b paşi (adesea cu mult mai puţini, deoarece diferenţa dintre două resturi succesive este de obicei mai mare decît 1), trebuie să apară restul 0:

$$r_{n-2} = r_{n-1}\, q_n + r_n$$
$$r_{n-1} = r_n\, q_{n+1} + 0.$$

Cînd se întîmplă acest lucru ştim că

$$(a, b) = r_n;$$

cu alte cuvinte (a, b) *este ultimul rest pozitiv din şirul* (5). Aceasta rezultă din aplicarea succesivă a egalităţii (3) egalităţilor (4), deoarece din (4) deducem

$$(a, b) = (b, r_1), \quad (b, r_1) = (r_1, r_2), \quad (r_1, r_2) = (r_2, r_3),$$
$$(r_2, r_3) = (r_3, r_4), \dots, (r_{n-1}, r_n) = (r_n, 0) = r_n.$$

Exerciţii: Efectuaţi algoritmul lui Euclid pentru găsirea celui mai mare divizor comun al lui (a) 187 şi 77, (b) 105 şi 385, (c) 245 şi 193.

Din ecuaţiile (4), poate fi dedusă o proprietate foarte importantă a lui (a, b). Dacă $d = (a, b)$, *atunci există întregi pozitivi sau negativi* k şi l, astfel încît

$$(6) \qquad d = ka + lb.$$

Pentru a arăta aceasta, să considerăm şirul (5) de resturi succesive. Din prima ecuaţie din (4)

$$r_1 = a - q_1b,$$

astfel încît r_1 se poate scrie sub forma $k_1a + l_1b$. (În acest caz $k_1 = 1$, $l_1 = -q_1$). Din egalitatea următoare obținem

$$r_2 = b - q_2r_1 = b - q_2(k_1a + l_1b) = (-q_2k_1)a + (1 - q_2l_1)b = \\ = k_2a + l_2b.$$

Este clar că acest procedeu poate fi continuat pînă ce ajungem la reprezentarea

$$r_n = ka + lb,$$

pe care am vrut s-o obținem.

Ca exemplu să considerăm algoritmul lui Euclid pentru găsirea lui (61, 24): cel mai mare divizor comun este 1 și reprezentarea care ne interesează pentru 1 se obține din egalitățile

$$61 = 2 \cdot 24 + 13, \quad 24 = 1 \cdot 13 + 11, \quad 13 = 1 \cdot 11 + 2,$$
$$11 = 5 \cdot 2 + 1, \quad 2 = 2 \cdot 1 + 0.$$

Din prima din aceste egalități

$$13 = 61 - 2 \cdot 24,$$

din a doua,

$$11 = 24 - 13 = 24 - (61 - 2 \cdot 24) = -61 + 3 \cdot 24,$$

din a treia,

$$2 = 13 - 11 = (61 - 2 \cdot 24) - (-61 + 3 \cdot 24) = 2 \cdot 61 - 5 \cdot 24,$$

și, în sfîrșit, din a patra

$$1 = 11 - 5 \cdot 2 = (-61 + 3 \cdot 24) - 5(2 \cdot 61 - 5 \cdot 24) = \\ = -11 \cdot 61 + 28 \cdot 24.$$

2. Aplicație la teorema fundamentală a aritmeticii

Faptul că $d = (a, b)$ poate fi scris întotdeauna sub forma $a = ka + lb$ poate fi folosit pentru a da o demonstrație teoremei fundamentale a aritmeticii, independentă de demonstrația dată la p. 39. Mai întîi vom demonstra, ca lemă, corolarul de la p. 40 și apoi, din această lemă, vom deduce teorema fundamentală inversînd ordinea demonstrației precedente.

Lemă: Dacă un număr prim p *divide un produs* ab, *atunci* p *divide pe* a *sau pe* b.

Dacă un număr prim p nu divide întregul a, atunci $(a, p) = 1$, deoarece singurii divizori ai lui p sînt p și 1. Deci, putem găsi întregii k și l, astfel încît

$$1 = ka + lp.$$

Înmulțind ambii membri ai acestei egalități cu b, obținem

$$b = kab + lpb.$$

Deoarece p divide pe ab, putem scrie

$$ab = pr,$$

astfel încît

$$b = kpr + lpb = p(kr + lb),$$

de unde se vede că p divide pe b. Am arătat deci că dacă p divide pe ab, dar nu divide pe a, atunci trebuie să dividă pe b; prin urmare, în orice caz, p trebuie să dividă pe a sau pe b, dacă divide pe ab.

Extinderea la produse de mai mult de doi factori este imediată. De exemplu, dacă p divide pe abc, atunci, aplicînd de două ori lema, putem arăta că p trebuie să dividă cel puțin unul din întregii a, b, și c. Deoarece dacă p nu divide nici pe a, b, nici pe c, atunci nu poate divide pe ab și deci nu poate divide pe $(ab)c = abc$.

Exercițiu: Extinderea acestui raționament la produse de n întregi necesită folosirea explicită, sau implicită, a principiului inducției matematice. Efectuați raționamentul cu toate amănuntele.

Din acest rezultat se deduce imediat teorema fundamentală a aritmeticii. Să presupunem că avem două descompuneri ale numărului întreg pozitiv N în factori primi:

$$N = p_1p_2 \ldots p_r = q_1q_2 \ldots q_s.$$

Deoarece p_1 divide membrul stîng al acestei egalități, el trebuie să dividă și membrul drept și deci, în baza exercițiului precedent, trebuie să dividă unul din factorii q_k. Dar, q_k este un număr prim și de aceea p_1 trebuie să fie egal cu q_k. După ce am simplificat egalitatea cu acești factori egali, ne îndreptăm atenția asupra factorului p_2, și stabilim în același mod că el este egal cu unul din factorii rămași q_t. Simplificînd pe p_2 și pe q_t continuăm în mod analog cu $p_3, \ldots p_r$. La capătul acestui proces toți factorii p au fost simplificați și în membrul stîng a rămas doar 1. Nici un q nu poate rămîne în membrul drept, deoarece toți factorii q sînt mai mari decît unu. Deci, factorii p și q vor fi grupați în perechi, doi cîte doi egali, ceea ce arată că, cu excepția ordinei factorilor, cele două descompuneri sînt identice.

3. Funcţia φ a lui Euler. Din nou despre teorema lui Fermat

Se spune că două numere întregi a şi b sînt *relativ prime*, dacă cel mai mare divizor comun al lor este egal cu 1:

$$(a, b) = 1.$$

De exemplu, numerele 24 şi 35 sînt relativ prime, în timp ce 12 şi 18 nu sînt. *Dacă* a *şi* b *sînt relativ prime, atunci putem alege numerele întregi* k *şi* l, *astfel încît*

$$ka + lb = 1.$$

Aceasta rezultă din proprietatea lui (a, b), enunţată la p. 61.

Exerciţiu: Demonstraţi teorema: *dacă un întreg* r *divide un produs* ab, *iar* r *şi* a *sînt relativ prime, atunci* r *trebuie să dividă pe* b. (Indicaţie: dacă r şi a sînt relativ prime, atunci putem găsi întregii k şi l, astfel încît:

$$kr + la = 1.$$

După aceea, înmulţiţi ambii membri ai acestei ecuaţii cu b). Această teoremă include lema de la p. 62, ca un caz particular, deoarece un număr prim p este relativ prim cu un întreg a, dacă şi numai dacă p nu divide pe a.

Pentru orice întreg pozitiv n, fie $\varphi(n)$ *numărul întregilor cuprinşi între* 1 *şi* n, *relativ primi cu* n. Această funcţie $\varphi(n)$, introdusă pentru prima dată de Euler, este o funcţie din teoria numerelor, de mare importanţă. Valorile lui $\varphi(n)$ pentru primele cîteva valori ale lui n se calculează uşor:

$\varphi(1) = 1$ deoarece 1 este relativ prim cu 1,
$\varphi(2) = 1$ deoarece 1 este relativ prim cu 2,
$\varphi(3) = 2$ deoarece 1 şi 2 sînt relativ prime cu 3,
$\varphi(4) = 2$ deoarece 1 şi 3 sînt relativ prime cu 4,
$\varphi(5) = 4$ deoarece 1, 2, 3, 4 sînt relativ prime cu 5,
$\varphi(6) = 2$ deoarece 1, 5 sînt relativ prime cu 6,
$\varphi(7) = 6$ deoarece 1, 2, 3, 4, 5, 6 sînt relativ prime cu 7,
$\varphi(8) = 4$ deoarece 1, 3, 5, 7 sînt relativ prime cu 8,
$\varphi(9) = 6$ deoarece 1, 2, 4, 5, 7, 8 sînt relativ prime cu 9,
$\varphi(10) = 4$ deoarece 1, 3, 7, 9 sînt relativ prime cu 10 etc.

Observăm că $\varphi(p) = p - 1$, dacă p este un număr prim; într-adevăr, un număr prim p nu are alţi divizori în afară de el însuşi şi de 1, şi deci este relatv prim cu toţi întregii $1, 2, 3, \ldots, p - 1$. Dacă n este un număr compus, cu descompunerea în factori primi

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r},$$

unde numerele p reprezintă factori primi distincţi, fiecare ridicat la o anumită putere, atunci

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right)\cdot\left(1 - \frac{1}{p_2}\right)\ldots\left(1 - \frac{1}{p_r}\right)\cdot$$

De exemplu, din descompunerea $12 = 2^2 \cdot 3$ rezultă

$$\varphi(12) = 12\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right) = 12\left(\frac{1}{2}\right)\left(\frac{2}{3}\right) = 4,$$

care se poate verifica imediat. Demonstraţia teoremei enunţate este elementară, dar nu o vom da.

**Exerciţiu:* Folosind funcţia φ a lui Euler, generalizaţi teorema lui Fermat de la p. 53. Teorema generală are următorul enunţ: *Dacă* n *este un întreg oarecare, iar* a *este relativ prim cu* n, *atunci*

$$a^{\varphi(n)} \equiv 1 \pmod{n}.$$

4. Fracţii continue. Ecuaţii diofantice

Algoritmul lui Euclid pentru găsirea celui mai mare divizor comun a doi întregi duce imediat la o metodă importantă de reprezentare a cîtului a doi întregi, printr-o fracţie compusă.

De exemplu, aplicînd algoritmul lui Euclid numerelor 840 şi 611, obţinem următoarele egalităţi

$$840 = 1 \cdot 611 + 229, \qquad 611 = 2 \cdot 229 + 153,$$
$$229 = 1 \cdot 153 + 76, \qquad 153 = 2 \cdot 76 + 1,$$

care arată, printre altele, că $(840, 611) = 1$. Din aceste egalităţi putem obţine următoarele expresii:

$$\frac{840}{611} = 1 + \frac{229}{611} = 1 + \frac{1}{611/229},$$

$$\frac{611}{229} = 2 + \frac{153}{229} = 2 + \frac{1}{229/153},$$

$$\frac{229}{153} = 1 + \frac{76}{153} = 1 + \frac{1}{153/76},$$

$$\frac{153}{76} = 2 + \frac{1}{76}\cdot$$

Combinînd aceste egalități, obținem dezvoltarea numărului rațional 840/611, sub forma

$$\frac{840}{611} = 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{76}}}}.$$

O expresie de forma

$$(7) \qquad a = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots + \cfrac{1}{a_n}}},$$

unde numerele a sînt întregi pozitivi, se numește *fracție continuă*. Algoritmul lui Euclid dă o metodă de exprimare a oricărui număr rațional sub forma unei fracții continue.

Exercițiu: Găsiți dezvoltarea în fracție continuă a numerelor

$$\frac{2}{5}, \quad \frac{43}{30}, \quad \frac{169}{70}.$$

* Fracțiile continue sînt deosebit de importante în acea ramură a aritmeticii superioare, cunoscută sub numele de analiză diofantică. O *ecuație diofantică* este o ecuație algebrică, cu una sau mai multe necunoscute și cu coeficienți întregi, pentru care se caută soluțiile *întregi*. O astfel de ecuație poate să nu aibă soluții, poate avea un număr finit sau chiar un număr infinit de soluții. Cel mai simplu caz este ecuația diofantică *liniară* cu două necunoscute,

$$(8) \qquad ax + by = c,$$

unde a, b și c sînt întregi dați și se se cere aflarea soluțiilor întregi x, y. Soluția generală a unei ecuații de acestă formă poate fi găsită cu ajutorul algoritmului lui Euclid.

Înainte de toate acest algoritm ne permite să determinăm pe $d = (a, b)$; după aceea, printr-o alegere potrivită a întregilor k și l, avem

$$(9) \qquad ak + bl = d.$$

Deci ecuația (8) are soluția particulară $x = k$, $y = l$, în cazul în care $c = d$. În general, dacă c este un multiplu al lui d:

$$c = d \cdot q,$$

atunci din (9) obţinem

$$a(kq) + b(lq) = dq = c,$$

astfel încît ecuaţia (8) are soluţia particulară $x = y^* = kq$, $y = y^* = lq$. Reciproc, dacă ecuaţia (8) are o soluţie x, y pentru un număr c dat, atunci c trebuie să fie un multiplu de $d = (a, b)$; pentru că d divide şi pe a şi pe b, şi deci trebuie să dividă pe c. Am demonstrat, prin urmare, că ecuaţia (8) are o soluţie, dacă şi numai dacă c este un multiplu de (a, b).

Pentru a determina celelalte soluţii ale lui (8), observăm că dacă $x = x'$, $y = y'$ este o soluţie diferită de soluţia $x = x^*$, $y = y^*$, găsită mai sus cu ajutorul algoritmului lui Euclid, atunci $x = x' - x^*$, $y = y' - y^*$ este o soluţie a ecuaţiei „omogene"

$$ax + by = 0. \tag{10}$$

Într-adevăr, dacă

$$ax' + by' = c \text{ şi } ax^* + by^* = c,$$

atunci, scăzînd a doua ecuaţie din prima, găsim că

$$a(x' - x^*) + b(y' - y^*) = 0.$$

Acum, soluţia generală a ecuaţiei (10) este $x = rb/(a, b)$, $y = -ra/(a, b)$, unde r este un întreg oarecare. (Lăsăm demonstraţia pe seama cititorului, ca exerciţiu. Indicaţie: împărţiţi cu (a, b) şi folosiţi exerciţiul de la p. 64). Rezultă imediat că

$$x = x^* + rb/(a, b), \qquad y = y^* - ra/(a, b).$$

În concluzie, ecuaţia diofantică liniară $ax + by = c$, unde a, b şi c sînt întregi, are o soluţie în numere întregi, dacă şi numai dacă c este multiplu de (a, b). În acest caz, o soluţie particulară $x = x^*$, $y = y^*$ poate fi găsită cu ajutorul algoritmului lui Euclid, iar soluţia generală este de forma

$$x = x^* + rb/(a,b), \quad y = y^* - ra/(a,b),$$

unde r este un întreg oarecare.

Exemple: Ecuaţia $3x + 6y = 22$ nu are soluţii întregi, deoarece $(3, 6) = 3$, care nu divide pe 22.

Ecuaţia $7x + 11y = 13$ are soluţia particulară $x = -39$, $y = 26$, care se obţine precum urmează:

$$11 = 1 \cdot 7 + 4, \qquad 7 = 1 \cdot 4 + 3, \qquad 4 = 1 \cdot 3 + 1, \qquad (7, 11) = 1.$$
$$1 = 4 - 3 = 4 - (7 - 4) = 2 \cdot 4 - 7 = 2(11 - 7) - 7 = 2 \cdot 11 - 3 \cdot 7.$$

Deci

$$7 \cdot (-3) + 11(2) = 1,$$
$$7 \cdot (-39) + 11(26) = 13.$$

Celelalte soluţii sînt date de

$$x = -39 + 11r, \qquad y = 26 - 7r,$$

unde r este un întreg oarecare.

Exerciţiu: Rezolvaţi ecuaţiile diofantice (a) $3x - 4y = 29$; (b) $11x + 12y = 58$; (c) $153x - 34y = 51$.

Capitolul II

SISTEMUL DE NUMERE AL MATEMATICII

INTRODUCERE

În cele ce urmează trebuie să extindem foarte mult noţiunea iniţială de număr, legată la început de şirul natural, pentru a crea un instrument destul de puternic pentru nevoile practicii şi ale teoriei. Într-o evoluţie îndelungată şi plină de ezitări, zero, întregii negativi şi fracţiile au fost acceptate doar treptat cu aceeaşi îndreptăţire ca şi întregii pozitivi, iar în zilele noastre, regulile de calcul cu aceste numere sînt stăpînite de copiii de vîrstă şcolară. Dar pentru a obţine o deplină libertate în operaţiile algebrice trebuie să mergem şi mai departe şi să includem cantităţile iraţionale şi complexe în noţiunea de număr. Cu toate că aceste extinderi ale noţiunii de număr natural au fost utilizate secole de-a rîndul şi se află la baza întregii matematici moderne, doar în timpurile recente ele au fost puse pe o bază logică solidă. În acest capitol vom face o expunere a acestei dezvoltări.

§ 1. NUMERELE RAŢIONALE

1. Numerele raţionale ca instrument de măsurare

Întregii au apărut ca abstracţii în procesul numărării unor colecţii finite de obiecte. Însă în viaţa de toate zilele trebuie nu numai să *numărăm obiecte* individuale, dar să şi *măsurăm* cantităţi, ca de pildă lungimea, aria, greutatea şi timpul. Dacă vrem să operăm cu uşurinţă cu rezultatele măsurătorilor acestor cantităţi, care admit subdivizări oricît de fine, este necesar să extindem limitele aritmeticii dincolo de întregi. Primul pas este să *reducem problema măsurării la problema numărării*. În primul rînd alegem, în mod arbitrar, o *unitate de măsură* — metrul, centimetrul, ţolul, kilogramul, gramul sau secunda, după caz — căreia îi atribuim măsura 1. Apoi, socotim numărul unităţilor de acest fel, care intră în cantitatea ce trebuie măsurată.

O anumită cantitate de plumb poate cîntări exact 54 kg. În general însă, procesul numărării unităților nu se va face „exact", și cantitatea dată nu va putea fi măsurată exact cu ajutorul multiplilor întregi ai unității alese. Tot ceea ce putem spune este că ea se află între doi multipli succesivi ai acestei unități, de pildă, între 53 și 54 kg. Dacă se întîmplă acest lucru, facem pasul următor, introducînd noi subunități, obținute prin subdivizarea unității inițiale, într-un număr de n părți egale. În limbajul obișnuit, aceste noi subunități pot avea nume speciale; de exemplu, metrul este divizat în 100 de centimetri, kilogramul în 1 000 de grame, ora în 60 de minute, minutul în 60 de secunde etc. În simbolismul matematicii însă, o subunitate obținută prin împărțirea unității inițiale 1 în n părți egale se notează cu simbolul $1/n$; dacă o anumită cantitate conține exact m astfel de subunități, măsura ei se notează cu simbolul m/n. Acest simbol se numește *fracție* sau *raport* (uneori este notată $m : n$). Următorul pas decisiv a fost făcut numai după secole de căutări dificile: simbolul m/n a fost dezbrăcat de legătura lui concretă cu procesul de măsurare și de cantitățile măsurate, fiind considerat în schimb ca *număr* pur, o entitate în sine, avînd aceeași îndreptățire ca și numerele naturale. Dacă m și n sînt numere naturale, simbolul m/n se numește *număr rațional*.

Folosirea cuvîntului număr (care inițial însemna doar număr natural) pentru aceste noi simboluri este justificată de faptul că adunarea și înmulțirea acestor simboluri ascultă de aceleași legi, care guvernează operațiile cu numere naturale. Pentru a arăta aceasta, trebuie să definim mai întîi adunarea, înmulțirea și egalitatea numerelor raționale. După cum știe oricine, aceste definiții sînt:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd},$$

(1)

$$\frac{a}{a} = 1, \qquad \frac{a}{b} = \frac{c}{d} \text{ dacă } ad = bc,$$

pentru orice întregi a, b, c, d. De exemplu:

$$\frac{2}{3} + \frac{4}{5} = \frac{2 \cdot 5 + 3 \cdot 4}{3 \cdot 5} = \frac{10 + 12}{15} = \frac{22}{15}, \quad \frac{2}{3} \cdot \frac{4}{5} = \frac{2 \cdot 4}{3 \cdot 5} = \frac{8}{15},$$

$$\frac{3}{3} = 1, \qquad \frac{8}{12} = \frac{6}{9} = \frac{2}{3}\cdot$$

Tocmai aceste definiții ne sînt impuse, dacă vrem să folosim numerele raționale pentru a măsura lungimi, arii etc. Dar riguros vorbind, aceste reguli pentru adunare, înmulțire și egalitate, referitoare la simbolurile noastre,

sînt stabilite prin propria noastră definiție și nu ne sînt impuse de vreo necesitate prealabilă, cu excepția necontradicției și a utilității pentru aplicații. Pe baza definițiilor (1), putem arăta că *legile fundamentale ale aritmeticii numerelor naturale continuă să rămînă valabile în domeniul numerelor raționale :*

(2) $p + q = q + p$ (legea comutativității adunării),
$p + (q + r) = (p + q) + r$ (legea asociativității adunării),
$pq = qp$ (legea comutativității înmulțirii),
$p(qr) = (pq)r$ (legea asociativității înmulțirii),
$p(q + r) = pq + pr$ (legea distributivității).

De exemplu, demonstrarea legii comutativității adunării pentru fracții este dată de egalitățile

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} = \frac{cb + da}{db} = \frac{c}{d} + \frac{a}{b},$$

dintre care prima și ultima egalitate corespund definiției (1) pentru adunare, în timp ce cea din mijloc este o consecință a legii comutativității adunării și înmulțirii pentru numerele naturale. Cititorul poate verifica în același mod celelalte patru legi.

Pentru o înțelegere reală a acestor fapte, trebuie să subliniem încă o dată că numerele raționale sînt propria noastră creație și că regulile (1) sînt impuse după voia noastră. În mod capricios, am putea decreta o altă lege pentru adunare, ca de pildă $\frac{a}{b} + \frac{c}{d} = \frac{a + c}{b + d}$, ceea ce, în particular, ar da $\frac{1}{2} + \frac{1}{2} = \frac{2}{4}$, rezultat absurd din punctul de vedere al măsurării. Reguli de acest fel, deși admisibile din punct de vedere logic, ar face din aritmetica simbolurilor noastre un joc lipsit de sens. Jocul liber al intelectului este condus aici de necesitatea creării unui instrument potrivit pentru efectuarea măsurătorilor.

2. Necesitatea intrinsecă a numerelor raționale. Principiul generalizării

În afară de motivul „practic" pentru introducerea numerelor raționale, mai există unul, mai intrinsec și, din anumite puncte de vedere, mai necesar, pe care îl vom discuta acum în mod cu totul independent de raționamentul precedent. El are un caracter în întregime aritmetic și este tipic, ca tendință dominantă, în raționamentul matematic.

În aritmetica obişnuită a numerelor naturale, putem efectua întotdeauna cele două operaţii fundamentale, adunarea şi înmulţirea. Dar „operaţiile inverse" de *scădere* şi *împărţire* nu sînt posibile întotdeauna. Diferenţa $b - a$ a doi întregi a, b este întregul c, astfel încît $a + c = b$, adică este soluţia ecuaţiei $a + x = b$. Dar în domeniul numerelor naturale simbolul $b - a$ are sens numai dacă $b > a$, deoarece numai atunci ecuaţia $a + x = b$ are ca soluţie x un număr natural. Un pas important în direcţia renunţării la această restricţie, a fost făcut prin introducerea simbolului 0, punîndu-se $a - a = 0$. De importanţă şi mai mare a fost introducerea simbolurilor -1, -2, -3, ..., împreună cu definiţia

$$b - a = -(a - b)$$

în cazul în care $b < a$; după aceasta s-a stabilit că scăderea poate fi făcută fără nici o restricţie *în domeniul întregilor pozitivi şi negativi*. Pentru a include noile simboluri -1, -2, -3, ..., într-o aritmetică extinsă, care cuprinde atît numerele pozitive cît şi pe cele negative, trebuie desigur să definim operaţiile cu ele în aşa fel, încît *regulile iniţiale ale operaţiilor aritmetice să se păstreze*. De exemplu, regula

$$(-1)(-1) = 1, \tag{3}$$

care stă la baza înmulţirii întregilor negativi este o consecinţă a dorinţei noastre de a păstra legea distributivităţii $a(b + c) = ab + ac$. Într-adevăr, dacă am fi declarat că $(-1)(-1) = -1$, atunci, punînd $a = -1$, $b = 1$, $c = -1$, am fi avut $(-1)(1 - 1) = -1 - 1 = -2$, în timp ce, pe de altă parte, am fi avut $(-1)(1 - 1) = -1 \cdot 0 = 0$.

A fost necesar un timp îndelungat pentru ca matematicienii să înţeleagă că regula semnelor (3), împreună cu toate celelalte definiţii care guvernează fracţiile şi întregii negativi, nu pot fi „demonstrate". Ele sînt *create* de noi cu scopul de a obţine libertatea de operaţie, conservînd totuşi legile fundamentale ale aritmeticii. Ce *putem* — şi trebuie — să demonstrăm este doar faptul că pe baza acestor definiţii se conservă legile comutativităţii, asociativităţii şi distributivităţii, ale aritmeticii. Chiar marele Euler a recurs la un raţionament cu totul neconvingător pentru a arăta că $(-1)(-1)$ „trebuie" să fie egal cu $+1$. Pentru că, după cum a raţionat, produsul trebuie să fie sau $+1$, sau -1, şi nu poate fi egal cu -1, deoarece $-1 = (+1)(-1)$.

Tot aşa cum introducerea întregilor negativi şi a lui zero permite efectuarea fără restricţii a scăderii, introducerea numerelor fracţionare îndepărtează obstacolul analog din calea împărţirii. Cîtul $x = b/a$ a doi întregi a şi b, definit de ecuaţia

$$ax = b, \tag{4}$$

există *ca întreg* numai dacă a este un factor al lui b. Dacă nu se întîmplă acest lucru, ca de pildă atunci cînd $a = 2$, $b = 3$, introducem pur şi simplu un nou simbol b/a, pe care îl numim fracţie, supus regulii $a(b/a) = a$, astfel încît b/a este prin definiţie soluţia ecuaţiei (4). Inventarea fracţiilor ca noi simboluri numerice face posibilă împărţirea fără restricţie — cu excepţia *împărţirii cu zero*, care este *exclusă de la bun început*.

Expresii, ca de pildă 1/0, 3/0, 0/0 etc., vor fi pentru noi simboluri fără sens. Într-adevăr, dacă împărţirea cu 0 ar fi permisă, am putea deduce din egalitatea $0 \cdot 1 = 0 \cdot 2$ consecinţa absurdă $1 = 2$. Uneori este însă util să notăm astfel de expresii prin simbolul ∞ (a se citi „infinit"), *cu condiţia să nu încercăm să operăm cu simbolul ∞ ca şi cum ar fi supus regulilor obişnuite ale aritmeticii.*

Acum ne este clară semnificaţia pur aritmetică a sistemului *tuturor numerelor raţionale* — întregi şi fracţii, pozitive şi negative. În acest domeniu numeric extins nu numai că sînt valabile legile formale ale asociativităţii, comutativităţii şi distributivităţii, dar şi ecuaţiile $a + x = b$ şi $ax = b$ au întotdeauna soluţii, $x = b - a$ şi $x = b/a$, cu condiţia ca în ultimul caz să avem $a \neq 0$. Cu alte cuvinte, în domeniul numerelor raţionale, aşa-numitele *operaţii raţionale* — adunarea, scăderea, înmulţirea şi împărţirea — pot fi efectuate fără restricţie şi nu ne conduc niciodată în afara acestui domeniu. Un astfel de domeniu închis de numere se numeşte *cîmp*. Ne vom întîlni cu alte exemple de cîmpuri mai departe, în acest capitol, şi în capitolul III.

Extinderea unui domeniu prin introducerea unor noi simboluri, în aşa fel încît legile care sînt valabile în domeniul iniţial să rămînă valabile şi în domeniul extins, constituie un exemplu tipic, caracteristic principiului matematic de *generalizare*. Generalizarea de la numerele naturale la cele raţionale satisface atît necesitatea teoretică de îndepărtare a restricţiilor din calea scăderii şi împărţirii, cît şi necesitatea practică ca numerele să exprime rezultatele măsurătorilor. Tocmai datorită faptului că numerele raţionale îndeplinesc această dublă necesitate, ele sînt atît de importante. Aşa cum am văzut, această extindere a noţiunii de număr a devenit posibilă prin crearea unor noi numere, sub forma unor simboluri abstracte, ca de pildă 0, —2 şi 3/4. Astăzi, cînd operăm cu astfel de numere fără dificultăţi, este greu să credem că pînă în secolul al XVII-lea ele nu se bucurau de aceeaşi încredere ca şi întregii pozitivi şi că erau folosite, la nevoie, cu neîncredere şi înfiorare. Tendinţa umană inerentă de a adera la „concret", întruchipată în şirul numerelor naturale, a fost răspunzătoare de această încetineală cu care s-a produs pasul inevitabil. Numai în domeniul abstractului poate fi creat un sistem satisfăcător al aritmeticii.

3. Interpretarea geometrică a numerelor raționale

O interpretare geometrică edificatoare a sistemului numerelor raționale este dată de următoarea construcție.

Pe o dreaptă oarecare, „axa numerelor", marcăm segmentul de la 0 la 1 (fig. 8). Prin aceasta am stabilit lungimea segmentului unitate, pe care o putem alege după voie. Întregii pozitivi și negativi sînt reprezentați atunci



Fig. 8. Axa numerelor

printr-o mulțime de puncte echidistante pe axa numerelor, și anume numerele pozitive la dreapta punctului 0, iar cele negative la stînga lui. Pentru a reprezenta fracțiile cu numitorul n, împărțim fiecare din segmentele de lungime egală cu unitatea în n părți egale; punctele subdiviziunii reprezintă atunci fracțiile care au numitorul n. Dacă facem acest lucru pentru orice întreg n, toate numerele raționale vor fi reprezentate prin puncte pe axa numerelor. Aceste puncte le vom numi *puncte raționale* și vom folosi termenii „număr rațional" și „punct rațional" ca fiind sinonime.

În capitolul I, § 1 am definit relația $A < B$ pentru numerele naturale. Pe axa numerică această relație se exprimă astfel: dacă numărul natural A este mai mic decît numărul natural B, atunci punctul A se află la stînga punctului B. Deoarece relația geometrică are loc între *toate* punctele raționale, sîntem conduși la a încerca să extindem relația aritmetică, astfel încît să păstrăm ordinea geometrică relativă a punctelor corespunzătoare. Aceasta se face prin definiția următoare: se spune că numărul rațional A este *mai mic* decît numărul rațional $B(A < B)$, și că B este *mai mare* decît A $(B > A)$, dacă $B - A$ este pozitiv. Rezultă atunci că dacă $A < B$, punctele (numerele) cuprinse *între* A și B sînt acelea care sînt $> A$ și $< B$. Orice astfel de pereche de puncte distincte, împreună cu punctele dintre ele, se numește *segment* sau *interval*, și se notează $[A, B]$.

Distanța de la punctul A la origine, considerată ca fiind pozitivă, se numește *valoarea absolută* a lui A și este notată prin simbolul

$$|A|.$$

În cuvinte, dacă $A \geqslant 0$, avem $|A| = A$; dacă $A \leqslant 0$, avem $|A| = -A$. Este limpede că dacă A și B au același semn, egalitatea $|A + B| = |A| + |B|$ are loc, în timp ce, dacă A și B au semne diferite, avem $|A + B| < |A| + |B|$. Deci, combinînd aceste două afirmații, avem inegalitatea generală

$$|A + B| \leqslant |A| + |B|,$$

care este valabilă oricare ar fi semnele lui A și B.

Un fapt de importanță fundamentală este exprimat de propoziția: *Punctele raționale sînt dense pe dreaptă.* Prin aceasta înțelegem că în interiorul oricărui interval, oricît de mic, se află puncte raționale. Este suficient să luăm un numitor n suficient de mare, astfel încît intervalul $[0, 1/n]$ să fie mai mic decît intervalul $[A, B]$ considerat; atunci, cel puțin una din fracțiile m/n trebuie să se afle în interval. Prin urmare, nici un interval de pe dreaptă, oricît ar fi el de mic, nu este lipsit de puncte raționale. În plus, rezultă că trebuie să existe o infinitate de puncte raționale, în orice interval. Într-adevăr, dacă ar exista doar un număr finit, intervalul cuprins între oricare două puncte raționale consecutive ar fi lipsit de puncte raționale, ceea ce, după cum am văzut, nu este posibil.

§ 2. SEGMENTE INCOMENSURABILE, NUMERE IRAȚIONALE ȘI NOȚIUNEA DE LIMITĂ

1. Introducere

Comparînd mărimile a două segmente a și b, se poate întîmpla ca a să fie conținut în b, exact de un număr întreg r de ori. În acest caz, putem exprima măsura segmentului b, cu ajutorul măsurii lui a, spunînd că lungimea lui b este de r ori mai mare decît cea a lui a. Se poate întîmpla ca, deși nici un multiplu întreg al lui a să nu fie egal cu b, să putem totuși împărți pe a într-un număr n de segmente egale, fiecare de lungime a/n, astfel încît un multiplu întreg m al segmentului a/n să fie egal cu b:

$$b = \frac{m}{n} a. \tag{1}$$

În cazul cînd are loc o egalitate de forma (1), spunem că cele două segmente a și b sînt *comensurabile*, deoarece ele au ca măsură comună segmentul a/n, care se cuprinde de n ori în a și de m ori în b. Mulțimea tuturor segmentelor comensurabile cu a va fi mulțimea formată din segmentele a căror lungime

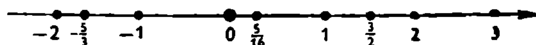


Fig. 9. Puncte raționale

poate fi exprimată sub forma (1), printr-o alegere corespunzătoare a întregilor m și n ($n \neq 0$). Dacă alegem segmentul a drept segment unitate $[0, 1]$, ca în fig. 9, atunci segmentele comensurabile cu segmentul unitate vor corespunde tuturor punctelor raționale m/n de pe axa numerelor. Pentru scopurile practice ale măsurării, numerele raționale sînt pe deplin satisfăcătoare. Chiar și din punct de vedere teoretic, deoarece mulțimea punctelor raționale

este densă pe dreaptă, s-ar putea să pară că toate punctele de pe dreaptă sînt puncte raționale. Dacă acest lucru ar fi adevărat, atunci orice segment ar fi comensurabil cu unitatea. Una dintre cele mai surprinzătoare descoperiri ale matematicii antice grecești (școala pitagoreică) a fost faptul că situația nu este cîtuși de puțin atît de simplă. Există *segmente incomensurabile* sau, dacă admitem că oricărui segment îi corespunde un număr care dă lungimea lui în raport cu unitatea, *numere iraționale*. Această revelație a constituit un eveniment științific de cea mai mare importanță. Este foarte posibil ca el să fi marcat originea raționamentului riguros în matematică, pe care o considerăm specifică matematicii grecești. Desigur, această descoperire a influențat profund întreaga matematică și chiar filozofia, din antichitate pînă în zilele noastre.

Teoria lui Eudoxus asupra mărimilor incomensurabile, prezentată sub formă geometrică în *Elementele* lui Euclid, este o capodoperă a matematicii grecești, cu toate că de obicei este omisă din versiunile diluate ale acestei opere clasice, folosite în școli. Teoria a fost pe deplin apreciată abia spre sfîrșitul secolului al XIX-lea, după ce Dedekind, Cantor și Weiersrtrass au dat o teorie riguroasă a numerelor iraționale. Vom prezenta această teorie sub aspectul aritmetic modern.

În primul rînd vom arăta că: *diagonala unui pătrat este incomensurabilă cu latura lui.* Să presupunem că latura pătratului este aleasă ca unitate de lungime și că diagonala are lungimea x. Atunci, din teorema lui Pitagora, avem

$$x^2 = 1^2 + 1^2 = 2.$$

(Putem nota pe x prin simbolul $\sqrt{2}$). Dacă x ar fi comensurabil cu 1, am putea găsi doi întregi p și q, astfel încît $x = p/q$ și

$$p^2 = 2q^2. \tag{2}$$

Se poate presupune că fracția p/q este ireductibilă, deoarece orice factor comun de la numărător și numitor ar putea fi simplificat de la început. Deoarece 2 apare ca factor în membrul drept, p^2 este un număr par și deci p este el însuși număr par, dat fiind că pătratul unui număr impar este impar. De aceea, putem scrie $p = 2r$. Ecuația (2) devine atunci

$$4r^2 = 2q^2 \quad \text{sau} \quad 2r^2 = q^2.$$

Deoarece 2 este un factor al membrului stîng, q^2 și deci și q trebuie să fie par. Astfel, p și q sînt ambele divizibile cu 2, ceea ce contrazice ipoteza că p și q nu au factor comun. Prin urmare, ecuația (2) nu poate avea loc și deci x nu poate fi număr rațional.

În alt mod, acest rezultat se poate formula prin propoziția că nu există nici un număr rațional egal cu $\sqrt{2}$.

Raționamentul din paragraful precedent arată că printr-o construcție geometrică foarte simplă se poate obține un segment incomensurabil cu unitatea. Dacă purtăm un astfel de segment pe axa numerelor cu ajutorul unui compas, punctul construit în acest fel nu poate coincide cu nici unul din punctele raționale: *sistemul punctelor raționale*, cu toate că este dens, *nu acoperă*

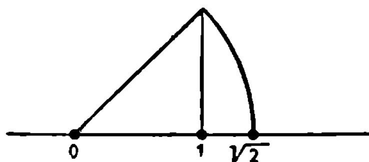


Fig. 10. Construcția lui $\sqrt{2}$

axa numerelor în întregime. Pentru o minte naivă, ar putea desigur să pară foarte straniu și paradoxal faptul că mulțimea densă a punctelor raționale nu acoperă dreapta în întregime. Nimic din „intuiția” noastră nu ne poate ajuta să „vedem” punctele iraționale sau să le deosebim de cele raționale. Nu este de mirare faptul că descoperirea incomensurabilului a tulburat matematicienii și filozofii greci și că pînă astăzi și-a păstrat efectul provocator asupra minților profunde.

Ar fi foarte ușor să construim oricît de multe segmente incomensurabile cu unitatea. Extremitățile acestor segmente, purtate pe axa numerică începînd din punctul 0, se numesc *puncte iraționale*. După cum principiul conducător în introducerea fracțiilor a fost *măsurarea lungimilor prin numere*, am vrea să menținem acest principiu la considerarea segmentelor incomensurabile cu unitatea. Dacă cerem să existe o *corespondență reciprocă între numere*, pe de o parte, și *punctele de pe o dreaptă*, pe de altă parte, este necesar să introducem *numerele iraționale*.

Recapitulînd situația de pînă acum, putem spune că un număr irațional reprezintă lungimea unui segment incomensurabil cu unitatea. În secțiunile următoare, vom rafina această definiție geometrică oarecum vagă, pînă ce vom ajunge la una mai satisfăcătoare din punctul de vedere al rigurozității logice. Primul pas în această direcție va fi făcut pe calea fracțiilor zecimale.

Exerciții: 1) Arătați că $\sqrt[3]{2}$, $\sqrt{3}$, $\sqrt{5}$, $\sqrt[3]{3}$ nu sînt raționale. (Indicație: folosiți lema de la p. 62.)

2) Arătați că $\sqrt{2}+\sqrt{3}$ și $\sqrt{2}+\sqrt[3]{2}$ nu sînt raționale. (Indicație: dacă, de exemplu, primul număr ar fi un număr rațional r, atunci, scriind $\sqrt{3}=r-\sqrt{2}$ și ridicînd la pătrat, am găsi că $\sqrt{2}$ ar fi rațional.)

3) Arătați că $\sqrt{2}+\sqrt{3}+\sqrt{5}$ este irațional. Încercați să găsiți exemple asemănătoare mai generale.

2. Fracții zecimale. Fracții zecimale infinite

Pentru a acoperi axa numerelor cu o mulțime de puncte peste tot densă, nu este necesar să folosim *toate* numerele raționale; este suficient, de exemplu, să considerăm numai acele numere care rezultă prin subdivizarea fiecărui interval unitate în 10, apoi în 100, 1 000 etc., segmente egale. Punctele obținute în acest fel corespund „fracțiilor zecimale". De exemplu, punctul $0{,}12 = 1/10 + 2/100$ corespunde punctului care se află în primul interval unitate, în al doilea subinterval de lungime 10^{-1}, și în punctul inițial al celui de-al treilea „sub-sub"— interval de lungime 10^{-2} (a^{-n} înseamnă $1/a^n$). Dacă o astfel de *fracție zecimală* conține n cifre după virgulă, atunci ea este de forma

$$f = z + a_1 10^{-1} + a_2 10^{-2} + a_3 10^{-3} + \ldots + a_n\, 10^{-n},$$

unde z este un întreg, iar coeficienții sînt cifre —0, 1, 2, ..., 9 — ce indică zecimile, sutimile ş.a.m.d. Numărul f este reprezentat în sistemul zecimal prin simbolul prescurtat $z, a_1a_2a_3 \ldots a_n$. Se vede imediat că aceste fracții zecimale pot fi scrise sub forma obişnuită a unei fracții p/q, în care $q = 10^n$; de exemplu, $f = 1{,}314 = 1 + 3/10 + 1/100 + 4/1000 = 1314/1000$. Dacă p şi q au un divizor comun, fracția zecimală poate fi redusă la o fracție al cărei numitor este un divizor al lui 10^n. Pe de altă parte, nici o fracție ireductibilă, al cărei numitor nu este un divizor al unei puteri a lui 10, nu poate fi reprezentată ca fracție zecimală. De exemplu, $\frac{1}{5} = \frac{2}{10} = 0{,}2$, şi $\frac{1}{250} = \frac{4}{1000} = 0{,}004$; dar $\frac{1}{3}$ nu poate fi scris ca fracție zecimală cu un număr finit n de cifre zecimale, oricît de mare ar fi n, pentru că o ecuație de forma

$$\frac{1}{3} = b/10^n$$

ar implica

$$10^n = 3b,$$

ceea ce este absurd, deoarece 3 nu este factor al nici unei puteri a lui 10.

Să alegem acum un punct P pe axa numerelor, care nu corespunde unei fracții zecimale; de exemplu, punctul rațional $\frac{1}{3}$ sau punctul irațional $\sqrt{2}$. Atunci, în procesul subdivizării intervalului unitate în zece părți egale ş.a.m.d., P nu va fi niciodată printre punctele de diviziune. Totuşi, P poate fi inclus în intervale din ce în ce mai mici ale diviziunii zecimale, cu orice grad de aproximare dorit. Procesul de aproximare poate fi descris precum urmează.

Să presupunem că P se află în primul interval unitate. Subdivizăm acest interval în zece părți egale, fiecare de lungime 10^{-1} și găsim, de pildă, că P se află în al treilea interval. La acest stadiu, putem spune că P se află între fracțiile zecimale 0,2 și 0,3. Subdivizăm intervalul cuprins între 0,2 și 0,3 în zece părți egale, fiecare de lungime 10^{-2}, și găsim că P se află, de pildă, în cel de-al patrulea interval. Subdivizîndu-l la rîndul său, găsim că P se află în primul interval de lungime 10^{-3}. Putem spune acum că P se află între 0,230 și 0,231. Procesul poate fi continuat la nesfîrșit și duce la un șir infinit de zecimale $a_1, a_2, a_3, \ldots, a_n, \ldots$, cu proprietatea următoare: oricare ar fi numărul n ales, punctul P este cuprins în intervalul I_n, a cărui extremitate din stînga este fracția zecimală $0, a_1a_2a_3 \ldots a_{n-1}a_n$ și a cărui extremitate din dreapta este $0, a_1a_2a_3 \ldots a_{n-1}(a_n + 1)$, lungimea lui I_n fiind egală cu 10^{-n}. Dacă alegem pe rînd $n = 1, 2, 3, 4, \ldots$, vedem că fiecare din aceste intervale $I_1, I_2, I_3, \ldots$, este conținut în cel precedent, în timp ce lungimile lor $10^{-1}, 10^{-2}, 10^{-3}, \ldots$ tind către zero. Spunem că punctul P este conținut într-un *șir descrescător de intervale zecimale*. De exemplu, dacă P este punctul rațional $\frac{1}{3}$, atunci toate zecimalele $a_1, a_2, a_3, \ldots$ sînt egale cu 3 și P este conținut în orice interval I_n, cuprins între 0,333 ... 33 și 0,333 ... 34; adică, $\frac{1}{3}$ este mai mare decît 0,333 ... 33, dar mai mic decît 0,333 ... 34, unde numărul zecimalelor poate fi luat oricît de mare. Exprimăm acest fapt spunînd că fracția zecimală cu n zecimale 0,333 ... 33 „tinde către $\frac{1}{3}$”, pe măsură ce n crește. Scriem

$$\frac{1}{3} = 0,333 \ldots,$$

punctele indicînd faptul că fracția zecimală trebuie continuată „la nesfîrșit”.

Punctul irațional $\sqrt{2}$ definit în secțiunea 1 duce, de asemenea, la o fracție zecimală infinită. Dar legea care determină valorile cifrelor din șir de data aceasta nu este cîtuși de puțin evidentă. De fapt, nu se cunoaște nici o formulă explicită, care să determine cifrele succesive, cu toate că putem calcula oricîte cifre dorim:

$$\begin{aligned} 1^2 = 1 < 2 < 2^2 &= 4 \\ (1,4)^2 = 1,96 < 2 < (1,5)^2 &= 2,25 \\ (1,41)^2 = 1,9881 < 2 < (1,42)^2 &= 2,0264 \\ (1,414)^2 = 1,999396 < 2 < (1,415)^2 &= 2,002225 \\ (1,4142)^2 = 1,99996164 < 2 < (1,4143)^2 &= 2,00024449 \text{ etc.} \end{aligned}$$

Ca definiție generală, spunem că un punct P, care nu poate fi reprezeutat sub forma unei fracții zecimale cu un număr finit n de zecimale, esereprezentat sub forma unei *fracții zecimale infinite* $z, a_1a_2a_3 \ldots$, dacă, pentru orice valoare a lui n, punctul P se află în intervalul de lungime 10^{-n}, care începe în punctul $z, a_1a_2a_3 \ldots a_n$.

În acest fel se stabilește o corespondență între toate punctele de pe axa numerelor și toate fracțiile zecimale *finite sau infinite*. Încercăm să dăm următoarea definiție: un „număr" este o fracție zecimală *finită sau infinită*. Acele fracții zecimale care nu reprezintă numere raționale se numesc numere *iraționale*.

Pînă la mijlocul secolului al XIX-lea, aceste considerații erau acceptate ca explicații satisfăcătoare ale sistemului de numere raționale și iraționale, *continuul numeric*. Progresul enorm făcut de matematică, începînd cu cel de-al XVII-lea secol, în particular dezvoltarea geometriei analitice și a calculului diferențial și integral, a fost posibil pe baza acestei reprezentări a sistemului de numere. Însă în perioada reexaminării critice a principiilor și a consolidării rezultatelor, s-a simțit din ce în ce mai mult necesitatea ca noțiunea de număr irațional să fie supusă unei analize mai profunde. Ca un preliminar al relatării noastre asupra teoriei moderne a continuului numeric vom discuta, într-un mod mai mult sau mai puțin intuitiv, noțiunea fundamentală de *limită*.

Exercițiu: Calculați $\sqrt[3]{2}$ și $\sqrt[3]{5}$ cu o aproximație de cel puțin 10^{-2}.

3. Limite. Serii geometrice infinite

După cum am văzut în secțiunea precedentă, se întîmplă uneori ca un anumit număr rațional s să fie aproximat printr-un șir de alte numere raționale s_n, unde indicele n ia succesiv toate valorile 1, 2, 3, ... De exemplu, dacă $s = 1/3$, atunci $s_1 = 0{,}3$, $s_2 = 0{,}33$, $s_3 = 0{,}333$ etc. Iată încă un exemplu. Să împărțim intervalul unitate în două jumătăți, a doua jumătate să o împărțim din nou în două părți egale, a doua din acestea din nou în două părți egale și așa mai departe, pînă ce cel mai mic din intervalele astfel obținute are lungimea 2^{-n}, unde n este un număr oricît de mare, de pildă, $n = 100$, $n = 100\,000$ sau orice număr dorim. Atunci, adunînd toate intervalele, cu excepția ultimului, obținem lungimea totală

$$s_n = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots + \frac{1}{2^n}. \tag{3}$$

Se vede ușor că s_n diferă de 1 prin $\left(\frac{1}{2}\right)^n$ și că această diferență devine oricît de mică sau „tinde către zero", pe măsură ce n crește nelimitat. Nu are

nici un sens să spunem că diferența *este egală* cu zero, dacă *n este* infinit. Infinitul intervine în matematică legat numai de un *proces* nelimitat și nu ca o *cantitate* reală. Descriem comportarea lui s_n, spunînd că *suma* s_n *se apropie de limita* 1, *pe măsură ce* n *tinde către infinit*, și vom scrie acest lucru în felul următor:

$$(4) \qquad 1 = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \cdots,$$

unde în membrul drept avem o *serie infinită*. Această „egalitate" nu înseamnă că avem de adunat într-adevăr o infinitate de termeni; ea este doar o expresie prescurtată a faptului că 1 este limita sumei finite s_n, cînd *n tinde* spre infinit (dar nu *este* nicidecum egal cu infinit). Astfel, egalitatea (4), cu simbolul ei incomplet „$+ \ldots$" este doar o prescurtare matematică pentru afirmația precisă

1 = limita, cînd *n* tinde către infinit, a cantității

$$(5) \qquad s_n = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^n}.$$

Sub o formă încă mai prescurtată, dar mai expresivă, scriem

$$(6) \qquad s_n \to 1 \text{ cînd } n \to \infty.$$

Vorbind despre limite, să mai dăm un exemplu. Să considerăm puterile unui număr q. Dacă $-1 < q < 1$, de pildă, $q = 1/3$ sau $q = -4/5$, atunci puterile succesive ale lui q

$$q, q^2, q^3, q^4, \ldots, q^n, \ldots,$$

tind către zero, pe măsură ce n crește. Dacă q este negativ, semnul lui q^n va alterna între $+$ și $-$, iar q^n va tinde către zero din ambele părți. Astfel, dacă $q = 1/3$, atunci $q^2 = 1/9$, $q^3 = 1/27$, $q^4 = 1/81$, ..., în timp ce, dacă $q = -1/2$, atunci $q^2 = 1/4$, $q^3 = -1/8$, $q^4 = 1/16$, ... Spunem că *limita lui* q^n, *cînd* n *tinde către infinit*, *este egală cu zero*, sau, simbolic:

$$(7) \qquad q^n \to 0 \text{ cînd } n \to \infty, \text{ dacă } -1 < q < 1.$$

(În treacăt fie zis, dacă $q > 1$ sau $q < -1$, atunci q^n nu tinde către zero, ci crește nelimitat în valoare absolută.)

Pentru a da o demonstrație riguroasă a afirmației (7), începem cu inegalitatea demonstrată la p. 32, care afirmă că $(1 + p)^n \geqslant 1 + np$ pentru orice întreg pozitiv n și $p > -1$. Fie q un număr fixat, cuprins între 0 și 1, de pildă, $q = 9/10$; avem $q = 1/(1 + p)$, unde $p > 0$. Deci

$$\frac{1}{q^n} = (1 + p)^n \geqslant 1 + np > np,$$

au (a se vedea regula 4, p. 340)

$$0 < q^n < \frac{1}{p} \cdot \frac{1}{n} \cdot$$

Rezultă că q^n este cuprins între numărul fix 0 şi numărul $(1/p)(1/n)$ care tinde spre zero pe măsură ce n creşte, deoarece p este fixat. Aceasta face evident faptul că $q^n \to 0$. Dacă q este negativ, avem $q = -1/(1+p)$ şi atunci q^n va fi cuprins între numerele $(-1/p)(1/n)$ şi $(1/p)(1/n)$ în loc de 0 şi $(1/p)(1/n)$. În rest, raţionamentul rămîne neschimbat.

Să considerăm acum *seria geometrică*

(8) $$s_n = 1 + q + q^2 + q^3 + \ldots + q^n.$$

(Cazul $q = 1/2$ a fost discutat mai sus). După cum s-a arătat, la p. 29), putem exprima suma s_n sub o formă simplă şi concisă. Dacă înmulţim pe s_n prin q, găsim

(8a) $$qs_n = q + q^2 + q^3 + q^4 + \ldots + q^{n+1},$$

şi scăzînd pe (8a) din (8), observăm că toţi termenii, cu excepţia lui 1 şi a lui q^{n+1}, se reduc. Prin acest artificiu obţinem

$$(1 - q)s_n = 1 - q^{n+1},$$

sau, prin împărţire,

$$s_n = \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q} - \frac{q^{n+1}}{1 - q}.$$

Noţiunea de limită apare dacă facem pe n să crească nelimitat. După cum am văzut, $q^{n+1} = q \cdot q^n$ tinde spre zero dacă $-1 < q < 1$, şi de aici putem conchide

(9) $$s_n \to \frac{1}{1 - q} \text{ cînd } n \to \infty, \text{ dacă } -1 < q < 1.$$

Acelaşi rezultat se poate scrie sub forma unei *serii geometrice infinite*

(10) $$1 + q + q^2 + q^3 + \ldots = \frac{1}{1 - q}, \text{ pentru } -1 < q < 1.$$

De exemplu,

$$1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots = \frac{1}{1 - \frac{1}{2}} = 2,$$

în concordanță cu egalitatea (4) și, în mod similar,

$$\frac{9}{10}+\frac{9}{10^2}+\frac{9}{10^3}+\frac{9}{10^4}+\ldots=\frac{9}{10}\cdot\frac{1}{1-1/10}=1,$$

astfel încît 0,99999... = 1. În mod analog, fracția zecimală finită 0,2374 și cea infinită 0,23739999999... reprezintă unul și același număr.

În capitolul VI vom reveni la discutarea generală a noțiunii de limită pe care o vom trata în spiritul riguros modern.

Exerciții: 1) Demonstrați că $1 - q + q^2 - q^3 + q^4 - \ldots = 1/(1+q)$, dacă $|q| < 1$.

2) Care este limita șirului $a_1, a_2, a_3, \ldots$, unde $a_n = n/(n+1)$? (Indicație: scrieți expresia sub forma $n/(n+1) = 1 - 1/(n+1)$ și observați că al doilea termen tinde către zero.)

3) Care este limita expresiei $\dfrac{n^2+n+1}{n^2-n+1}$ cînd $n \to \infty$?

(Indicație: Scrieți expresia sub forma $\dfrac{1+\frac{1}{n}+\frac{1}{n^2}}{1-\frac{1}{n}+\frac{1}{n^2}}$).

4) Demonstrați că dacă $|q| < 1$, atunci $1 + 2q + 3q^2 + 4q^3 + \ldots = \dfrac{1}{(1-q)^2}$. (Indicație: folosiți rezultatul exercițiului 3 de la p. 34.)

5) Care este limita seriei infinite

$$1 - 2q + 3q^2 - 4q^3 + \ldots ?$$

6) Care este limita expresiilor

$$\frac{1+2+3+\ldots+n}{n^2},\quad \frac{1^2+2^2+\ldots+n^2}{n^3} \text{ și } \frac{1^3+2^3+\ldots+n^3}{n^4}?$$

(Indicație: folosiți rezultatele de la pp. 28—31.)

4. Numerele raționale și fracțiile zecimale periodice

Numerele raționale p/q, care nu sînt fracții zecimale finite, pot fi dezvoltate sub forma unor fracții zecimale infinite, prin efectuarea procesului elementar al împărțirii. La fiecare stadiu al acestui proces trebuie să rămînă un rest nenul, deoarece în caz contrar, fracția zecimală ar fi finită. Toate resturile care apar în procesul împărțirii vor fi întregi cuprinși între 1 și $q - 1$, astfel încît sînt cel mult $q - 1$ posibilități diferite pentru valorile resturilor. Aceasta înseamnă că după cel mult q împărțiri, un rest k va apărea a doua oară. Dar atunci, toate resturile care urmează se vor repeta în aceeași ordine în care ele au apărut după prima apariție a restului k. Aceasta

arată că *reprezentarea zecimală a oricărui număr rațional este periodică*; după ce, la început, a apărut o mulțime finită de zecimale, aceeași zecimală sau grup de zecimale se vor repeta de o infinitate de ori. De exemplu, $1/6 = 0,1666666666...$; $1/7 = 0,142857142857142857...$; $1/11 = 0,09090909...$; $122/1100 = 0,1109090909...$; $11/90 = 0,122222222...$ etc. (Numerele raționale care pot fi reprezentate sub forma unor fracții zecimale finite pot fi gîndite ca fracții zecimale periodice, în care cifra 0 se repetă de o infinitate de ori, după un număr finit de zecimale.) Din exemplele date se vede că unele fracții zecimale periodice au o parte neperiodică, care precede partea periodică.

Reciproc, se poate arăta că toate *fracțiile zecimale periodice sînt numere raționale*. Să considerăm ca exemplu fracția zecimală periodică infinită

$$p = 0,3322222...$$

Avem $p = 33/100 + 10^{-3}\,2(1 + 10^{-1} + 10^{-2} + ...)$. Expresia din paranteze este seria geometrică infinită

$$1 + 10^{-1} + 10^{-2} + 10^{-3} + \cdots = \frac{1}{1 - 1/10} = \frac{10}{9}.$$

Deci

$$p = \frac{33}{100} + 2 \cdot 10^{-3} \cdot \frac{10}{9} = \frac{2970 + 20}{9 \cdot 10^3} = \frac{2990}{9000} = \frac{299}{900}.$$

Demonstrația pentru cazul general este în esență aceeași, dar pretinde o notație mai generală. În fracția zecimală periodică generală

$$p = 0, a_1a_2a_3 \ldots a_mb_1b_2 \ldots b_nb_1b_2 \ldots b_nb_1b_2 \ldots b_n \ldots$$

punem $0, b_1b_2 \ldots b_n = B$, astfel încît B reprezintă partea periodică a fracției zecimale. Atunci p devine

$$p = 0, a_1a_2 \ldots a_m + 10^{-m} B(1 + 10^{-n} + 10^{-2n} + 10^{-3n} \ldots).$$

Expresia din paranteze este o serie geometrică infinită, în care $q = 10^{-n}$. Suma ei, potrivit egalității (10) din secțiunea precedentă, este egală cu $1/(1 - 10^{-n})$ și de aceea

$$p = 0, a_1a_2 \ldots a_m + \frac{10^{-m} B}{1 - 10^{-n}}.$$

Exerciții: 1) Dezvoltați fracțiile $\frac{1}{11}$, $\frac{1}{13}$, $\frac{2}{13}$, $\frac{3}{13}$, $\frac{1}{17}$, $\frac{2}{17}$, în fracții zecimale și determinați perioada.

*2) Numărul 142 857 are proprietatea că prin înmulțire cu oricare din numerele 2, 3, 4, 5, sau 6 se produce o permutare ciclică a cifrelor sale. Explicați această proprietate folosind dezvoltarea lui 1/7 în fracție zecimală.

3) Dezvoltați numerele raționale din exercițiul 1) ca fracții „zecimale", cu bazele 5, 7 și 12.

4) Dezvoltați pe 1/3 sub forma unei fracții zecimale.

5) Scrieți numărul 0,11212121... sub forma unei fracții ordinare. Găsiți valoarea acestui simbol, dacă el este gîndit în sistemele cu bazele 3 sau 5.

5. Definiția generală a numerelor iraționale cu ajutorul șirurilor de intervale

La p. 79 am încercat să adoptăm o definiție: un „număr" este o fracție zecimală finită sau infinită. Am admis că acele fracții zecimale infinite, care nu reprezintă numere raționale, să fie numite numere iraționale. Pe baza rezultatelor din secțiunea precedentă, putem formula această definiție în modul următor: continuul numeric, sau *sistemul numerelor reale* („real" în contrast cu numerele „imaginare" sau „complexe", care vor fi introduse în § 5) este *totalitatea fracțiilor zecimale infinite* (fracțiile zecimale finite pot fi considerate drept caz particular, în care toate zecimalele, începînd dintr-un anumit loc, sînt egale cu zero sau, tot la fel de bine, am putea conveni ca în loc să luăm o fracție zecimală finită, a cărui ultimă zecimală este a, să scriem o fracție zecimală infinită, în care în locul lui a se află $a - 1$ urmat de un număr infinit de zecimale, toate egale cu 9. Aceasta exprimă faptul că 0,999... = 1, potrivit secțiunii 3). Numerele *raționale* sînt fracțiile zecimale *periodice*; numerele *iraționale* sînt fracțiile zecimale *neperiodice*. Chiar această definiție nu pare a fi pe deplin satisfăcătoare; pentru că, așa cum am văzut în cap. I, prin însăși natura lucrurilor sistemul zecimal nu este cu nimic privilegiat față de alte sisteme posibile. Tot atît de bine am fi putut face raționamentul cu sistemul diadic sau cu oricare altul. Din acest motiv, este de dorit să dăm o definiție mai generală a continuului numeric, independentă de vreo referință specială la baza 10. Poate că cel mai simplu mod de a face acest lucru este următorul:

Să considerăm un șir $I_1, I_2, \ldots, I_n, \ldots$ de intervale pe axa numerelor, cu extremități raționale, fiecare fiind conținut în precedentul și astfel încît lungimea celui de-al n-lea interval I_n să tindă către zero, cînd n crește. Un astfel de șir se numește *șir descrescător*[1] *de intervale*. În cazul intervalelor zecimale, lungimea lui I_n este 10^{-n}, dar tot atît de bine ar putea fi 2^{-n} sau ar putea fi restrînsă doar prin condiția ca ea să fie mai mică decît $1/n$. Acum formulăm ca postulat de bază al geometriei: *oricărui șir descrescător*

[1] În original, *sequence of nested intervals*, care s-ar traduce prin șir de intervale cuibărite. — *N.T.*

de intervale îi corespunde un singur punct pe axa numerelor, care este conținut în același timp în toate intervalele. (Se vede direct că nu poate exista mai mult de *un* punct comun tuturor intervalelor, pentru că lungimile intervalelor tind către zero, și două puncte diferite nu ar putea fi conținute în nici un interval mai mic decît distanța dintre ele.) Acest punct se numește, prin definiție, *număr real*; dacă el nu este un punct rațional, se numește *număr irațional.* Prin această definiție stabilim o corespondență perfectă între puncte și numere. Aceasta nu este decît o formulare mai generală a celor exprimate cu ajutorul definiției care folosea fracții zecimale infinite.

Aici cititorul ar putea fi cuprins de o îndoială pe deplin legitimă. Ce *este* acest „punct" de pe axa numerelor, pe care îl presupunem comun tuturor intervalelor unui șir descrescător de intervale, în cazul în care nu este un punct rațional? Răspunsul nostru este următorul: existența pe axa numerelor (privită ca dreaptă) a unui punct conținut în orice șir descrescător de intervale, cu extremități raționale, este un *postulat* fundamental *al geometriei.*

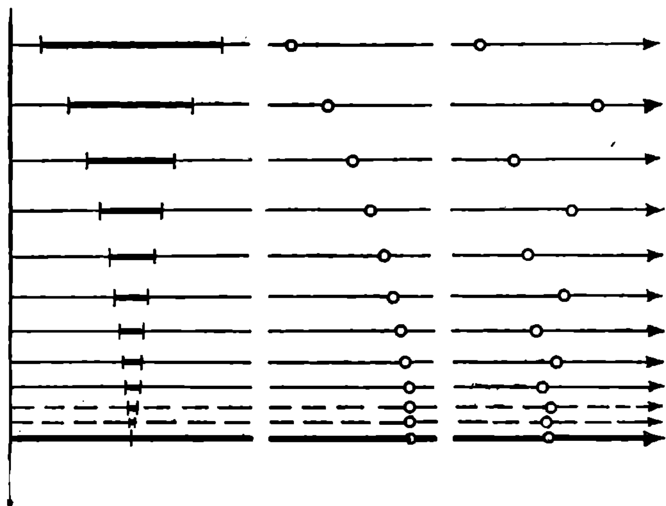


Fig. 11. Șir descrescător de intervale. Limite de șiruri

Nu se cere nici o reducere logică a acestui postulat la alte fapte matematice. Îl acceptăm, tot așa cum acceptăm alte axiome sau postulate în matematică, deoarece este plauzibil din punct de vedere intuitiv și este util pentru construirea unui sistem consistent de gîndire matematică. Din punct de vedere pur formal, am putea începe cu o dreaptă formată numai din puncte raționale și apoi am putea *defini* punctul irațional, ca fiind tocmai un *simbol*

Cazul în care A are un cel mai mare element a^* şi B are un cel mai mic element b^* este imposibil, deoarece atunci numărul raţional $(a^* + b^*)/2$, care se află la mijloc, între a^* şi b^*, ar fi mai mare decît cel mai mare element al lui A şi mai mic decît cel mai mic element al lui B, şi deci n-ar putea aparţine nici uneia din aceste clase.

În cel de-al treilea caz, în care nu există nici cel mai mare număr raţional în A şi nici cel mai mic număr raţional în B, Dedekind spune că tăietura defineşte, sau pur şi simplu *este*, un număr iraţional. Se vede cu uşurinţă că această definiţie este în concordanţă cu definiţia dată cu ajutorul şirului descrescător de intervale; orice şir descrescător $I_1, I_2, I_3, \ldots$ de intervale defineşte o tăietură, dacă punem în clasa A toate acele numere raţionale care sînt depăşite de cel puţin una din extremităţile iniţiale ale intervalelor I_n, iar în B toate celelalte numere raţionale.

Din punct de vedere filozofic, definiţia lui Dedekind a numerelor iraţionale implică un grad de abstracţie destul de înalt, deoarece nu impune nici o restricţie asupra naturii legii matematice care defineşte cele două clase A şi B. O metodă mai concretă de definire a continuului numerelor reale aparţine lui Georg Cantor (1845—1918). Cu toate că la prima vedere este cu totul diferită de metoda şirurilor descrescătoare de intervale sau de cea a tăieturilor, ea este echivalentă cu fiecare din ele, în sensul că sistemele de numere definite pe aceste trei căi au aceleaşi proprietăţi. Ideea lui Cantor a fost sugerată de următoarele fapte: 1) numerele reale pot fi privite ca fracţii zecimale infinite şi 2) fracţiile zecimale infinite sînt limite de fracţii zecimale finite. Eliberîndu-ne de dependenţa de sistemul zecimal, vom spune, ca şi Cantor, că orice şir convergent $a_1, a_2, a_3, \ldots$ de numere raţionale defineşte un *număr real*. Prin convergenţă se înţelege faptul că diferenţa $(a_m - a_n)$ dintre doi termeni ai şirului tinde către zero, dacă a_m şi a_n sînt destul de departe în şir, adică dacă m şi n tind spre infinit. (Aproximaţiile zecimale succesive ale oricărui număr au această proprietate, deoarece oricare două, după cea de-a n-a, diferă prin cel mult 10^{-n}.) Deoarece există multe căi prin care putem aproxima acelaşi număr real printr-un şir de numere raţionale, spunem că două şiruri convergente de numere raţionale $a_1, a_2, a_3, \ldots$ şi $b_1, b_2, b_3, \ldots$ definesc acelaşi număr real, dacă $a_n - b_n$ tinde către zero, cînd n creşte indefinit. Operaţiile de adunare etc., pentru astfel de şiruri sînt foarte uşor de definit.

§ 3. OBSERVAŢII ASUPRA GEOMETRIEI ANALITICE[2]

1. Principiul fundamental

Continuul numeric, fie că este acceptat de la început ca un lucru de la sine înţeles, fie numai după o examinare critică, a devenit baza matematicii şi, în particular, a geometriei analitice şi calculului diferenţial şi integral, încă din secolul al XVII-lea.

[2] Pentru cititorii care nu sînt familiarizaţi cu subiectul acestui paragraf sînt date cîteva exerciţii în apendicele de la sfîrşitul cărţii, pp. 509—514.

Introducînd continuul numeric, devine posibilă asocierea unui număr real determinat fiecărui segment de dreaptă, prin care se definește lungimea lui. Dar se poate merge și mai departe. Nu numai lungimea, dar *orice obiect geometric și orice operație geometrică pot fi transpuse în domeniul numerelor.* Pașii decisivi spre această aritmetizare a geometriei au fost făcuți, încă din 1629, de Fermat (1601—1655), și în 1637, de Descartes (1596—1650). Ideea fundamentală a geometriei analitice este introducerea „coordonatelor", adică a *numerelor* atașate sau coordonate unui *obiect geometric*, prin care acesta este complet caracterizat. Majoritatea cititorilor cunosc așa-numitele coordonate rectangulare sau carteziene, care servesc la caracterizarea poziției unui punct arbitrar în plan. Începem cu două drepte perpendiculare, fixe, din plan, „axa x-ilor" și „axa y-ilor", față de care raportăm orice punct. Aceste drepte sînt considerate ca fiind axe numerice orientate și înzestrate cu aceeași unitate de măsură. Fiecărui punct P (fig. 12) i se asociază două coordonate x și y. Acestea se obțin în modul următor: considerăm segmentul orientat, care pornește din „originea" O și se termină în punctul P „vectorul de poziție" al punctului P, și apoi proiectînd acest segment orientat pe cele două axe, obținem segmentul orientat OP' pe axa x-lor a cărui lungime este dată de numărul x și, de asemena, segmentul orientat OQ', de pe axa y-ilor, a cărui lungime este dată de numărul y. Cele două numere x și y se numesc *coordonatele* lui P. Reciproc, dacă x și y sînt două numere

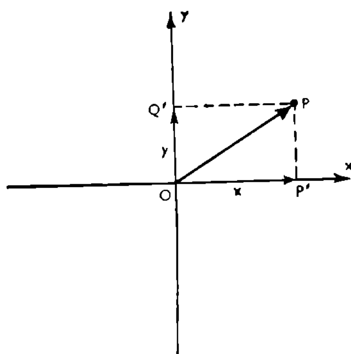


Fig. 12. Coordonate rectangulare ale unui punct

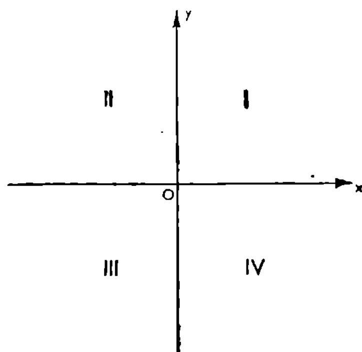


Fig. 13. Cele patru cadrane

date arbitrar, atunci punctul corespunzător P este determinat în mod unic. Dacă x și y sînt ambele pozitive, P se află în *primul cadran* al sistemului de coordonate (fig. 13); dacă ambele sînt negative, P se află în cel de-al treilea cadran; dacă x este pozitiv, iar y negativ, el se află în cel de-al patrulea, iar dacă x este negativ și y pozitiv, în cel de-al doilea cadran.

Distanța dintre punctul P_1, de coordonate x_1, y_1 și punctul P_2 de coordonate x_2, y_2, este dată de formula

$$d^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2. \tag{1}$$

Aceasta rezultă imediat din teorema lui Pitagora, după cum se poate vedea din fig. 14.

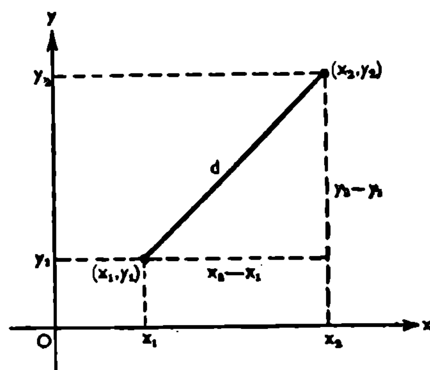


Fig. 14. Distanța dintre două puncte

*2. Ecuațiile dreptelor și curbelor

Dacă C este un punct fixat de coordonate $x = a$, $y = b$, atunci locul tuturor punctelor P care se află la o distanță dată r de C este un cerc care are punctul C ca centru și raza r. Din formula (1) a distanței rezultă că punctele acestui cerc au coordonate x, y, care satisfac ecuația

$$(x - a)^2 + (y - b)^2 = r^2. \tag{2}$$

Aceasta se numește *ecuația cercului*, pentru că ea exprimă condiția completă (necesară și suficientă) impusă coordonatelor x, y ale punctului P, pentru ca el să se afle pe cercul cu centrul C și de rază r. Dacă desfacem parantezele, ecuația (2) devine

$$x^2 + y^2 - 2ax - 2by = k, \tag{3}$$

unde $k = r^2 - a^2 - b^2$. Reciproc, dacă este dată o ecuație de forma (3), unde a, b și k sînt constante arbitrare, astfel încît $k + a^2 + b^2$ să fie un număr pozitiv, atunci prin procedeul algebric de „completare a pătratului" putem scrie ecuația sub forma:

$$(x - a)^2 + (y - b)^2 = r^2,$$

unde $r^2 = k + a^2 + b^2$. Rezultă că ecuația (3) definește un cerc de rază r, cu centrul C, ale cărui coordonate sînt a și b. Ecuațiile dreptelor au o formă și mai simplă. De exemplu, axa x-lor are ecuația $y = 0$, deoarece $y = 0$ pentru toate punctele de pe axa x-lor și numai pentru acestea. Axa y-lor are ecuația $x = 0$. Dreptele duse prin origine, bisectoare ale unghiuri-

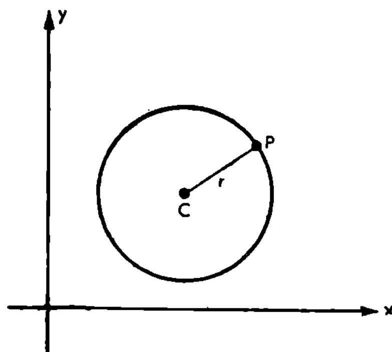


Fig. 15. Cercul

lor formate de axe, au ecuațiile $x = y$ și $x = -y$. Se vede cu ușurință că orice dreaptă are o ecuație de forma

$$(4) \qquad ax + by = c,$$

unde a, b, c sînt constante fixe, care caracterizează dreapta. Înțelesul ecuației (4) constă din nou în faptul că toate perechile de numere reale, care satisfac această ecuație sînt coordonatele unui punct al dreptei și reciproc.

S-ar putea ca cititorul să fi învățat că ecuația

$$(5) \qquad \frac{x^2}{p^2} + \frac{y^2}{q^2} = 1$$

reprezintă o elipsă (fig. 16). Această curbă taie axa x-lor în punctele $A(p, 0)$ și $A'(-p, 0)$, iar axa y-lor în $B(0, q)$ și $B'(0, -q)$. (Notația $P(x, y)$ sau pur și simplu (x, y) este folosită ca prescurtare în loc de „punctul P de coordonate x și y"). Dacă $p > q$, segmentul AA', de lungime $2p$, se numește axa mare a elipsei, în timp ce segmentul BB', de lungime $2q$, se numește axa mică. Această elipsă este locul geometric al tuturor punctelor P, pentru care suma distanțelor la punctele $F(\sqrt{p^2 - q^2}, 0)$ și $F'(-\sqrt{p^2 - q^2}, 0)$ este egală cu $2p$. Ca exercițiu, cititorul poate verifica acest lucru, folosind

formula (1). Punctele F și F' se numesc *focarele* elipsei, iar raportul $e = \frac{\sqrt{p^2 - q^2}}{p}$ se numește *excentricitatea* elipsei.

O ecuație de forma

$$\frac{x^2}{p^2} - \frac{y^2}{q^2} = 1 \tag{6}$$

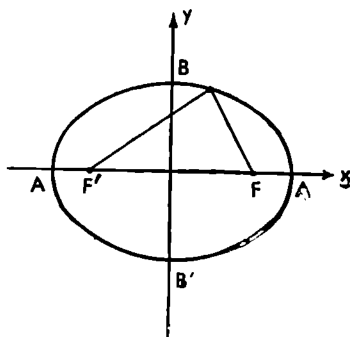


Fig. 16. Elipsa; F *și* F' *sînt focarele*

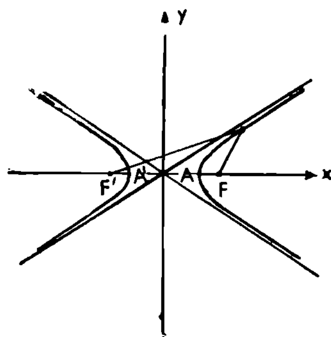


Fig. 17. Hiperbola; F *și* F' *sînt focarele*

reprezintă o hiperbolă. Ea este formată din două ramuri care taie axa x-lor respectiv în $A(p, 0)$ și $A'(-p, 0)$ (fig. 17). Segmentul AA', de lungime $2p$, se numește axa transversă a hiperbolei. Hiperbola se apropie din ce în ce mai mult de dreptele $qx \pm py = 0$, pe măsură ce ne îndepărtăm de origine, dar nu atinge niciodată aceste drepte. Ele se numesc *asimptotele* hiperbolei. Hiperbola este locul geometric al tuturor punctelor P, pentru care diferența distanțelor la punctele $F(\sqrt{p^2 + q^2}, 0)$ și $F'(-\sqrt{p^2 + q^2}, 0)$ este egală cu $2p$. Aceste puncte se numesc, în cazul hiperbolei, tot focare; prin excentricitate se înțelege raportul $e = \frac{\sqrt{p^2 + q^2}}{p}$.

Ecuația

$$xy = 1 \tag{7}$$

definește de asemenea o hiperbolă, ale cărui asimptote sînt cele două axe (fig. 18). Ecuația acestei hiperbole „ echilatere" indică faptul că aria dreptunghiului determinat de P este egală cu 1, oricare ar fi punctul P de pe curbă. O hiperbolă echilateră, a cărei ecuație este

$$xy = c, \tag{7a}$$

unde c este constantă, este un caz particular al hiperbolei generale, tot aşa cum cercul este un caz particular al elipsei. Caracteristica specială a hiperbolei echilatere constă în faptul că cele două asimptote (în cazul nostru cele două axe de coordonate) sînt perpendiculare.

Pentru noi este importantă aici ideea fundamentală că obiectele geometrice pot fi reprezentate complet într-o formă aritmetică sau algebrică. Acelaşi lucru este adevărat şi despre operaţiile geometrice. De exemplu, dacă vrem să găsim punctul de intersecţie a două drepte, considerăm ecuaţiile lor

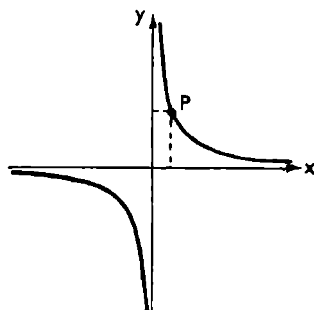


Fig. 18. Hiperbola echilaterală xy = 1. *Aria dreptunghiului determinat de punctul* P(x, y) *este egală cu 1.*

$$ax + by = c \tag{8}$$
$$a'x + b'y = c'.$$

Punctul comun celor două drepte se obţine atunci, pur şi simplu, prin determinarea coordonatelor lui ca soluţie x, y a sistemului de ecuaţii (8). În mod analog, punctele de intersecţie a două curbe oarecare, ca de pildă cercul $x^2 + y^2 - 2ax - 2by = k$ şi dreapta $ax + by = c$, pot fi găsite prin rezolvarea sistemului corespunzător de ecuaţii.

§ 4. ANALIZA MATEMATICĂ A INFINITULUI

1. Noţiuni fundamentale

Şirul întregilor pozitivi

$$1, 2, 3, \ldots$$

este primul şi cel mai important exemplu de mulţime infinită. Nu este nici un mister în faptul că acest şir nu are sfîrşit, nu se termină; într-adevăr,

oricît de mare ar fi întregul *n*, întregul următor, $n + 1$, poate fi format întotdeauna. Dar trecînd de la *adjectivul* „infinit", ceea ce înseamnă pur şi simplu „fără sfîrşit", la *substantivul* „infinit", nu trebuie în nici un caz să facem ipoteza că „infinitul", exprimat de obicei prin simbolul special ∞, poate fi considerat ca şi cum ar fi un *număr* obişnuit. Nu putem include simbolul ∞ în sistemul numerelor reale, conservînd în acelaşi timp regulile fundamentale ale aritmeticii. Cu toate acestea, noţiunea de infinit inundă întreaga matematică, deoarece obiectele matematice sînt studiate de obicei, nu ca indivizi izolaţi, ci ca membri ai unor clase sau sisteme care conţin o infinitate de obiecte de acelaşi tip; de pildă, totalitatea întregilor, a numerelor reale, a triunghiurilor din plan. Pentru acest motiv este necesar să analizăm infinitul matematic într-un mod precis. Teoria modernă a mulţimilor, creată la sfîrşitul secolului al XIX-lea de Georg Cantor şi de şcoala sa, a răspuns acestei necesităţi cu un succes răsunător. Teoria mulţimilor a lui Cantor a pătruns şi a influenţat puternic multe domenii ale matematicii şi a devenit de importanţă fundamentală în studiul fundamentelor logice şi filozofice ale matematicii. Punctul de plecare este noţiunea generală de *mulţime* sau *sistem*. Prin aceasta se înţelege orice colecţie de obiecte, definite printr-o regulă, care specifică în mod exact care sînt obiectele care aparţin colecţiei date. Ca exemple, putem considera mulţimea tuturor întregilor pozitivi, mulţimea tuturor fracţiilor zecimale periodice, mulţimea tuturor numerelor reale, sau mulţimea tuturor dreptelor din spaţiul tridimensional.

Pentru a compara „mărimea" a două mulţimi diferite se foloseşte noţiunea fundamentală de „echivalenţă". Dacă elementele din două mulţimi A şi B pot fi grupate în perechi, în aşa fel încît fiecărui element a lui A să-i corespundă un singur element din B, iar fiecărui element din B să-i corespundă un singur element din A, atunci se spune că corespondenţa este *biunivocă*, iar mulţimile A şi B se numesc *echivalente*. Noţiunea de echivalenţă pentru mulţimile *finite* coincide cu noţiunea obişnuită de *egalitate a numerelor*, deoarece două mulţimi finite au acelaşi număr de elemente dacă şi numai dacă elementele celor două mulţimi pot fi puse în corespondenţă biunivocă. Aceasta este de fapt chiar ideea numărării, deoarece atunci cînd numărăm o mulţime finită de obiecte stabilim, pur şi simplu, o corespondenţă biunivocă între aceste obiecte şi o mulţime de simboluri numerice $1, 2, 3, \ldots, n$.

Pentru a stabili echivalenţa a două mulţimi finite, nu este întotdeauna necesar să numărăm obiectele ei. De exemplu, putem afirma, fără a număra, că orice mulţime finită de cercuri de rază 1 este echivalentă cu mulţimea centrelor lor.

Ideea lui Cantor a fost să extindă noţiunea de echivalenţă la mulţimile infinite cu scopul de a defini o „artimetică" a infiniţilor. Mulţimea tuturor numerelor reale este echivalentă cu mulţimea tuturor punctelor de pe o

dreaptă, deoarece alegerea unei origini și a unei unități de lungime ne permite să-i asociem, în mod biunivoc, oricărui punct P de pe dreaptă, un anumit număr real x drept coordonată:

$$P \longleftrightarrow x.$$

Întregii pari formează o submulțime proprie a mulțimii *tuturor întregilor*, iar *întregii* formează o submulțime proprie a mulțimii *numerelor raționale* (prin termenul *submulțime proprie* a unei mulțimi S, înțelegem o mulțime S' care constă din unele dar nu din toate elementele din S). Desigur, *dacă o mulțime este finită*, adică dacă ea conține un număr de n elemente și nu mai multe, *atunci ea nu poate fi echivalentă cu nici una din submulțimile ei proprii*, deoarece orice submulțime proprie ar putea conține cel mult $n - 1$ elemente. Dar, *dacă o mulțime conține o infinitate de elemente*, atunci, ceea ce este destul de paradoxal, *ea poate fi echivalentă cu o submulțime proprie a ei*. De exemplu, corespondența

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & & n & \ldots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \\ 2 & 4 & 6 & 8 & 10 & \ldots & 2n & \ldots \end{array}$$

este o corespondență biunivocă între mulțimea *întregilor pozitivi* și submulțimea proprie a *întregilor pari*, care se arată a fi echivalente. Acest fapt, care contrazice adevărul familiar „întregul este mai mare decît oricare din părțile lui", arată la ce fel de surprize ne putem aștepta în domeniul aritmeticii infinitului.

2. Numărabilitatea numerelor raționale și nenumărabilitatea continuului

Una din primele descoperiri ale lui Cantor în analiza pe care a făcut-o asupra infinitului a fost aceea că mulțimea *numerelor raționale* (care conține mulțimea infinită a întregilor și este deci ea însăși infinită) este echivalentă cu *mulțimea întregilor*. La prima vedere pare foarte straniu ca mulțimea densă a numerelor raționale să fie la fel de numeroasă ca și submulțimea rară a întregilor. Desigur, nu putem așeza numerele raționale pozitive în *ordinea mărimii* (cum putem face cu întregii) spunînd că a este primul număr rațional, b al doilea ca mărime și așa mai departe, pentru că există o infinitate de numere raționale între oricare două și deci nu există un număr „următor". Dar, așa cum a observat Cantor, neglijînd relația de mărime dintre două elemente succesive, este posibil să așezăm toate numerele raționale într-un singur rînd $r_1, r_2, r_3, r_4, \ldots$, așa cum putem face cu întregii. În acest șir va exista un prim număr rațional, un al doilea, al treilea și așa mai departe, și fiecare număr rațional va apare o singură dată. O ast-

fel de aranjare a unei mulțimi de obiecte într-un șir, ca acela al întregilor, se numește *numărare* a mulțimii. Indicînd o astfel de numărare, Cantor a arătat că mulțimea numerelor raționale este echivalentă cu mulțimea întregilor, deoarece corespondența

$$\begin{array}{cccccccc} 1 & 2 & 3 & 4 & \ldots & n & \ldots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \\ r_1 & r_2 & r_3 & r_4 & \ldots & r_n & \ldots \end{array}$$

este biunivocă. O metodă de numărare a numerelor raționale va fi descrisă în cele ce urmează.

Orice număr rațional poate fi scris sub forma a/b, unde a și b sînt întregi, și toate aceste numere pot fi așezate într-un tablou în care a/b se află în coloana a și în linia b. De exemplu, 3/4 se află în cea de-a treia coloană și cea de-a patra linie a tabloului de mai jos (fig. 19). Toate numerele raționale pozitive pot fi așezate acum conform schemei următoare: în tabloul din fig. 19 trasăm o linie frîntă continuă, care trece prin toate numerele tabloului. Începînd cu 1, facem un pas spre dreapta și obținem 2 drept al doilea termen al șirului, apoi mergînd pe diagonală spre stînga în jos obținem 1/2 următorul pas pe verticală în jos ne dă 1/3, după aceea pe diagonală în sus,

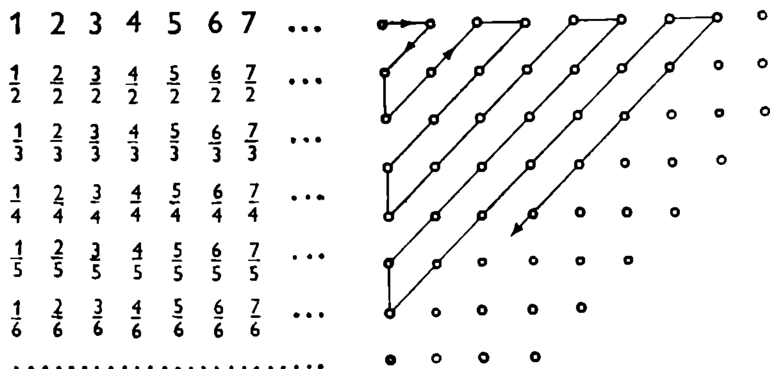


Fig. 19. Numărarea numerelor raționale

la dreapta, pînă ce atingem din nou prima linie în 3, apoi la 4, pe diagonală în jos la 1/4 și așa mai departe, după cum se arată în figură. Mergînd în lungul acestei linii frînte, ajungem la un șir 1, 2, 1/2, 1/3, 2/2, 3, 4, 3/2, 2/3, 1/4, 1/5, 2/4, 3/3, 4/2, 5, ... care conține numerele raționale în ordinea în care ele se succed de-a lungul liniei frînte. Dacă ștergem din acest șir toate numerele a/b, în care a și b au un divizor comun, obținem un șir în care fiecare număr rațional r apare exact o dată; 1, 2, 1/2, 1/3, 3, 4,

3/2, 2/3, 1/4, 1/5, 5, ... Aceasta arată că mulțimea tuturor numerelor raționale pozitive este numărabilă. Datorită faptului că numerele raționale corespund într-un mod biunivoc punctelor raționale de pe o dreaptă, am demonstrat în același timp că mulțimea punctelor raționale pozitive de pe o dreaptă este numărabilă.

Exerciții: 1) Arătați că mulțimea tuturor întregilor pozitivi și negativi este numărabilă. Arătați că mulțimea tuturor numerelor raționale pozitive și negative este numărabilă.

2) Arătați că mulțimea $S + T$ (cf. p. 127) este numărabilă dacă S și T sînt mulțimi numărabile. Arătați același lucru pentru suma a trei, patru sau a oricărui număr n de mulțimi și, în fîrșit, pentru o mulțime formată dintr-o mulțime numărabilă de mulțimi numărabile.

Deoarece am arătat că numerele raționale sînt numărabile, am putea bănui că orice mulțime infinită este numărabilă și că prin aceasta, evident, s-ar termina întreaga analiză a infinitului. Lucrurile nu stau cîtuși de puțin astfel. Cantor a făcut descoperirea foarte importantă că *mulțimea tuturor numerelor reale*, raționale și iraționale, *nu este numărabilă*. Cu alte cuvinte, totalitatea numerelor reale prezintă un tip de infinitate cu totul deosebit (cum se spune, superior) aceluia al întregilor sau al numerelor raționale. Demonstrația indirectă, ingenioasă, a lui Cantor a acestui fapt a devenit un model pentru multe demonstrații matematice. Schița demonstrației este următoarea: Pornim de la ipoteza că toate numerele reale au fost numărate într-un șir, iar apoi indicăm un număr care nu apare în numărarea presupusă. De aici rezultă o contradicție, deoarece ipoteza era că *toate* numerele reale au fost incluse în numărare, și această ipoteză trebuie să fie falsă, chiar dacă un singur număr a fost omis. Deci, ipoteza că este posibilă numărarea numerelor reale este inadmisibilă și deci negația ei, adică propoziția lui Cantor că mulțimea numerelor reale nu este numărabilă, este dovedită.

Pentru a efectua acest raționament să presupunem că toate numerele reale, prezentate sub forma unor fracții zecimale infinite, sînt așezate într-un tabel

primul număr	$N_1, a_1\, a_2\, a_3\, a_4\, a_5 \ldots$
al doilea număr	$N_2, b_1\, b_2\, b_3\, b_4\, b_5 \ldots$
al treilea număr	$N_3, c_1\, c_2\, c_3\, c_4\, c_5 \ldots$

. . . .

unde literele N reprezintă partea întreagă, iar literele minuscule desemnează cifrele aflate după virgulă. Presupunem că acest șir de fracții zecimale conține *toate* numerele reale. Punctul esențial al demonstrației constă acum în construirea unui nou număr, printr-un „proces diagonal”, despre care vom arăta că nu se află în acest șir. Să construim un astfel de număr. Pentru aceasta alegem mai întîi o cifră a, diferită de a_1 și de 0 sau 9 (pentru a evita ambiguitățile posibile, care pot apărea din egalități, ca de pildă

0,999 ... = 1,000 ...), apoi o cifră b diferită de b_2 şi de 0 sau 9, în mod analog c diferită de c_3 şi aşa mai departe (de exemplu, putem alege, pur şi simplu, $a = 1$ dacă $a_1 \neq 1$, şi $a = 2$ dacă $a_1 = 1$, şi în mod asemănător pentru celelalte numere din tabel $b, c, d, e, \ldots$). Să considerăm acum fracţia zecimală infinită

$$z = 0, a\ b\ c\ d\ e \ldots$$

Acest nou număr z este cu siguranţă diferit de oricare din numerele din tabelul de mai sus; el nu poate fi egal cu primul, deoarece diferă de el prin prima cifră după virgulă; nu poate fi egal cu al doilea, deoarece diferă prin a doua cifră după virgulă şi, în general, nu poate fi egal cu cel de-al n-lea număr din tabel, deoarece diferă prin cea de-a n-a cifră după virgulă. Aceasta arată că tabelul nostru de fracţii zecimale aşezate succesiv *nu* conţine toate numerele reale. Deci, această mulţime nu este numărabilă.

Cititorul şi-ar putea închipui că motivul nenumărabilităţii continuului numeric constă în faptul că dreapta este infinită şi că un segment finit al dreptei ar conţine doar o infinitate numărabilă de puncte. Lucrurile nu stau astfel, deoarece se poate arăta uşor că întregul continuu numeric este echivalent cu orice segment finit, ca de pildă segmentul cuprins între 0 şi 1 (fără

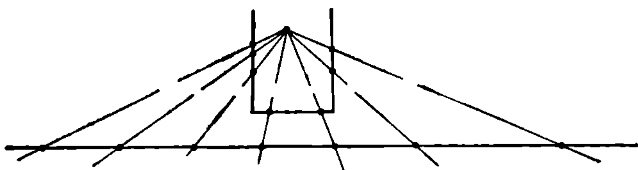


Fig. 20. Corespondenţa biunivocă dintre punctele unui segment îndoit şi punctele unei drepte

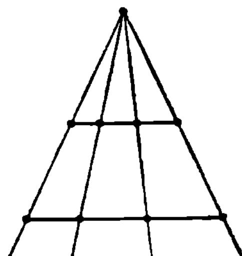


Fig. 21. Corespondenţa biunivocă dintre punctele a două segmente de lungime diferită

extremităţi). Corespondenţa biunivocă cerută poate fi obţinută îndoind segmentul în punctele 1/3 şi 2/3 şi proiectîndu-l apoi, aşa cum se arată în fig. 20. Rezultă deci că chiar un segment finit al axei numerice conţine o infinitate nenumărabilă de puncte.

Exerciţiu: Arătaţi că orice interval $[A, B]$ al axei numerice este echivalent cu oricare alt interval $[C, D]$.

Este folositor să indicăm o altă demonstrație, poate mai intuitivă, a nenumărabilității continuului numeric. Conform celor arătate va fi suficient să ne concentrăm atenția asupra mulțimii punctelor cuprinse între 0 și 1. Demonstrația este și de data aceasta indirectă. Să presupunem că mulțimea tuturor punctelor de pe dreaptă, cuprinse între 0 și 1, poate fi așezată într-un șir

$$(1) \qquad a_1, a_2, a_3, \ldots .$$

Să includem punctul de coordonată a_1 într-un interval de lungime $1/10$, punctul de coordonată a_2 într-un interval de lungime $1/10^2$ și așa mai departe. Dacă toate punctele cuprinse între 0 și 1 ar fi incluse în șirul (1), intervalul unitate ar fi acoperit în întregime de șirul infinit de astfel de subintervale, de lungimi $1/10$, $1/10^2$, ..., care eventual se suprapun. (Faptul că unele din ele depășesc intervalul unitate nu influențează demonstrația noastră.) Suma acestor lungimi este dată de seria geometrică

$$1/10 + 1/10^2 + 1/10^3 + \ldots = \frac{1}{10}\left[\frac{1}{1-\frac{1}{10}}\right] = \frac{1}{9} \cdot$$

Astfel, ipoteza că șirul (1) conține toate numerele reale cuprinse între 0 și 1 ne conduce la posibilitatea acoperirii complete a unui interval de lungime 1 printr-o mulțime de intervale de lungime totală $1/9$, ceea ce este absurd din punct de vedere intuitiv. Am putea accepta această contradicție ca o demonstrație, cu toate că, din punct de vedere logic, ea ar necesita o analiză mai aprofundată.

Raționamentul din paragraful precedent servește la stabilirea unei teoreme de mare importanță în teoria modernă a „măsurii". Înlocuind intervalele de mai sus prin intervale mai mici, de lungime $\varepsilon/10^n$, unde ε este un număr pozitiv oricît de mic, vedem că orice mulțime numărabilă de puncte de pe dreaptă poate fi inclusă într-o mulțime de intervale de lungime totală egală cu $\varepsilon/9$. Deoarece ε a fost arbitrar, ultimul număr poate fi făcut oricît de mic dorim. În terminologia teoriei măsurii spunem că o mulțime numărabilă de puncte are *măsura nulă*.

Exercițiu : Demonstrați că același rezultat rămîne valabil pentru o mulțime numărabilă de puncte din plan, înlocuind lungimile intervalelor prin ariile pătratelor.

3. „Numerele cardinale" ale lui Cantor

Să rezumăm rezultatele de pînă acum : numărul de elemente dintr-o mulțime *finită* A nu poate fi egal cu numărul de elemente dintr-o mulțime finită *B*, dacă *A* conține *mai multe* elemente decît *B*. Dacă înlocuim noțiunea de „mulțimi cu același număr (finit) de elemente" cu noțiunea mai generală de *mulțimi echivalente*, atunci pentru mulțimi infinite propoziția precedentă nu mai este adevărată : mulțimea tuturor întregilor conține mai multe ele-

mente decît mulțimea întregilor pari, iar mulțimea numerelor raționale mai multe decît mulțimea întregilor; și totuși am văzut că aceste mulțimi sînt echivalente. Am putea bănui că *toate* mulțimile infinite sînt echivalente și că alte distincții decît aceea dintre numerele finite și infinit nu pot fi făcute, dar rezultatul lui Cantor infirmă această presupunere; există o mulțime, continuul numerelor reale, care nu este echivalentă cu nici o mulțime numărabilă.

Există deci cel puțin două tipuri diferite de „infinit", infinitul numărabil al întregilor și infinitul nenumărabil al continuului. Dacă două mulțimi A și B, finite sau infinite, sînt echivalente, vom spune că ele au același *număr cardinal*. Dacă A și B sînt mulțimi finite, numărul cardinal se reduce la noțiunea obișnuită de număr natural, dar noțiunea de număr cardinal are un caracter mult mai general. Mai mult, dacă o mulțime A este echivalentă cu o submulțime a lui B, în timp ce B nu este echivalentă cu A sau cu nici una din submulțimile ei, vom spune, după Cantor, că mulțimea B are un *număr cardinal mai mare* decît acela al mulțimii A. Această utilizare a cuvîntului „număr" concordă cu noțiunea obișnuită de număr folosită în cazul mulțimilor finite. Mulțimea întregilor este o submulțime a mulțimii numerelor reale, în timp ce mulțimea numerelor reale nu este echivalentă nici cu mulțimea întregilor, nici cu oricare din submulțimile ei (adică, mulțimea numerelor reale nu este nici numărabilă, nici finită). Deci, potrivit definiției noastre, continuul numerelor reale are un număr cardinal mai mare decît acela al mulțimii întregilor.

De fapt, Cantor a arătat cum să construim un întreg șir de mulțimi infinite cu numere cardinale din ce în ce mai mari. Deoarece putem începe cu mulțimea întregilor pozitivi, este desigur suficient să arătăm că *fiind dată orice mulțime* A, *putem construi o altă mulțime* B *cu un număr cardinal mai mare*. Datorită generalității acestei teoreme, demonstrația este în mod necesar oarecum abstractă. Definim mulțimea B ca fiind mulțimea ale cărei elemente sînt toate submulțimile mulțimii A. Prin cuvîntul „submulțime" vom înțelege nu numai submulțimile proprii ale lui A, dar chiar și mulțimea A, și „submulțimea" vidă 0, care nu conține nici un element. (Dacă A constă din întregii 1, 2, 3, atunci B conține 8 elemente diferite $\{1, 2, 3\}$, $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, $\{1\}$, $\{2\}$, $\{3\}$, și 0.) Fiecare element al mulțimii B este el însuși o *mulțime* formată din anumite elemente ale lui A. Să presupunem acum că B este echivalentă cu A sau cu o submulțime a ei, adică că există o regulă care pune în corespondență, într-un mod biunivoc, elementele lui A, sau ale unei submulțimi a lui A, cu toate elementele lui B, adică cu submulțimile lui A:

$$a \leftrightarrow S_a, \tag{2}$$

unde am notat cu S_a submulțimea lui A, care corespunde elementului a al lui A. Vom ajunge la o contradicție indicînd un element al lui B (adică o submulțime T a lui A), care nu poate corespunde nici unui element a. Pentru a construi această submulțime, observăm că pentru orice element x al lui A există două posibilități: sau mulțimea S_x, asociată lui x prin cores-

pondența dată (2), conține elementul x, sau S_x nu conține pe x. *Definim pe* T *ca fiind submulțimea lui* A, *care constă din toate elementele* x *care au proprietatea că* S_x *nu îl conține pe* x. Această submulțime se deosebește de orice S_a prin cel puțin elementul a, deoarece dacă S_a conține pe a, T nu îl conține, în timp ce dacă S_a nu îl conține pe a, T îl conține. Deci T nu este inclus în corespondența (2). Aceasta arată că este imposibil de stabilit o corespondență biunivocă între elementele lui A sau ale oricărei submulțimi a lui A și acelea ale lui B. Însă corespondența

$$a \leftrightarrow \{a\}$$

este o corespondență biunivocă între elementele lui A și submulțimea lui B formată din toate submulțimile lui A, care constau dintr-un singur element. Deci, din definiția cuprinsă în paragraful precedent, B are un număr cardinal mai mare decît acela al lui A.

Exercițiu: Dacă A conține n elemente, unde n este un întreg pozitiv, arătați că B, definită ca mai sus, conține 2^n elemente. Dacă A constă din mulțimea tuturor întregilor pozitivi, arătați că B este echivalentă cu continuul numerelor reale, cuprinse între 0 și 1. (Indicație: simbolizați o submulțime a lui A, în primul caz, printr-un șir finit, iar în al doilea caz, printr-un șir infinit de simboluri 0 și 1,

$$a_1 a_2 a_3 \ldots,$$

unde $a_n = 1$ sau 0, după cum al n-lea element al lui A aparține sau nu aparține submulțimii date.)

Ar putea să pară că este simplu să găsim o mulțime de *puncte* cu un număr cardinal mai mare decît mulțimea numerelor reale cuprinse între 0 și 1. Desigur, un pătrat, fiind „bidimensional", pare să conțină „mai multe" puncte decît un segment „unidimensional". Destul de surprinzător, lucrurile nu stau așa: *numărul cardinal al mulțimii punctelor dintr-un pătrat este același ca și numărul cardinal al mulțimii punctelor de pe un segment.* Pentru a demonstra acest lucru, indicăm corespondența următoare.

Dacă (x, y) este un punct al pătratului unitate, x și y pot fi scrise sub formă zecimală în felul următor:

$$x = 0{,}a_1a_2a_3a_4 \ldots,$$

$$y = 0{,}b_1b_2b_3b_4 \ldots,$$

unde, pentru a evita ambiguitatea, am ales de pildă pe 0,250000... în locul lui 0,249999... pentru a reprezenta numărul rațional 1/4. Punctului (x, y) din pătrat îi asociem punctul

$$z = 0{,}a_1b_1a_2b_2a_3b_3a_4b_4 \ldots,$$

cuprins în segmentul de extremități 0 și 1. Desigur, puncte diferite (x, y) și (x', y') ale pătratului vor corespunde unor puncte diferite z și z' ale segmentului, astfel încît numărul cardinal al pătratului nu poate depăși pe acela al segmentului.

(De fapt, corespondența pe care tocmai am definit-o este biunivocă între mulțimea punctelor pătratului și o submulțime proprie a segmentului unitate; de exemplu, nici un punct al pătratului nu poate corespunde punctului 0,2140909090..., deoarece forma 0,25000... a fost

preferată formei 0,24999... pentru desemnarea numărului 1/4. Putem însă să modificăm puţin corespondenţa, astfel încît ea să devină biunivocă între întreg pătratul şi întreg segmentul, care au în acest fel acelaşi număr cardinal.)

Un raţionament analog arată că numărul cardinal al punctelor dintr-un cub nu este mai mare decît numărul cardinal al segmentului.

Cu toate că aceste rezultate par să contrazică noţiunea intuitivă de dimensiune, trebuie să ne reamintim că corespondenţa pe care am definit-o nu este „continuă"; dacă parcurgem în mod continuu segmentul de la 0 la 1, punctele corespunzătoare ale pătratului nu vor forma o curbă continuă, ci vor apare într-o ordine complet haotică. Dimensiunea unei mulţimi de puncte depinde nu numai de numărul cardinal al mulţimii, dar şi de modul în care punctele sînt distribuite în spaţiu. În capitolul V vom reveni asupra acestei probleme.

4. Metoda indirectă de demonstrare

Teoria numerelor cardinale reprezintă numai un aspect al teoriei generale a mulţimilor, creată de Cantor cu toate criticile severe ale unora dintre cei mai distinşi matematicieni ai timpului. Mulţi dintre aceşti critici ca, de pildă, Kronecker şi Poincaré, au obiectat faţă de imprecizia noţiunii generale de „mulţime" şi faţă de caracterul neconstructiv al raţionamentului folosit pentru definirea anumitor mulţimi.

Obiecţiile aduse raţionamentului neconstructiv se referă la ceea ce s-ar putea numi *demonstraţii esenţial indirecte*. Demonstraţiile indirecte sînt un fel familiar de raţionament matematic: pentru a stabili adevărul unei propoziţii A, încercăm să presupunem că A', negaţia lui A, este adevărată. Apoi, printr-un lanţ de raţionamente producem o contradicţie a lui A', demonstrînd în acest fel absurditatea lui A'. Deci, pe baza principiului logic fundamental al „terţiului exclus", absurditatea lui A' stabileşte adevărul lui A.

În diferite locuri din carte cititorul va întîlni un şir de exemple pentru care o demonstraţie indirectă poate fi înlocuită cu uşurinţă printr-o demonstraţie directă, cu toate că forma indirectă a demonstraţiei are avantajul conciziei şi absenţei unor detalii care nu sînt necesare pentru obiectivul imediat. Există totuşi teoreme pentru care nu s-au putut da decît demonstraţii indirecte. Există chiar teoreme demonstrabile prin metoda indirectă pentru care demonstraţiile constructive directe nu pot fi date nici măcar în principiu, tocmai datorită naturii acestor teoreme. Astfel, de exemplu, este teorema de la p. 97. Nu o dată în istoria matematicii s-a întîmplat ca atunci cînd eforturile matematicienilor erau îndreptate spre *construirea* soluţiilor anumitor probleme, pentru a demonstra rezolvabilitatea lor, altcineva să vină şi, ocolind problema construcţiei, să dea o demonstraţie indirectă, neconstructivă.

Există o diferenţă esenţială între demonstrarea existenţei unui obiect de un anumit tip prin construirea unui exemplu palpabil al obiectului şi a demon-

stra că, dacă nu ar exista nici unul, s-ar putea deduce rezultate contradictorii. În primul caz, avem un obiect palpabil, în timp ce în al doilea avem numai o contradicție. Unii matematicieni distinși au susținut eliminarea, mai mult sau mai puțin completă, a tuturor demonstrațiilor neconstructive din matematică. Chiar dacă un astfel de program ar fi de dorit, ar antrena în prezent o complicație imensă și chiar distrugerea parțială a corpului matematicii existente. Din acest motiv, nu este de mirare faptul că școala „intuiționismului", care a adoptat acest program, a întîmpinat o rezistență puternică și chiar cei mai fanatici intuiționiști nu pot trăi întotdeauna după convingerile lor.

5. Paradoxurile infinitului

Cu toate că poziția fără compromisuri a intuiționiștilor este prea extremistă pentru cei mai mulți matematicieni, o amenințare serioasă la adresa frumoasei teorii a sistemelor infinite a apărut atunci cînd în sînul ei s-au ivit adevărate paradoxuri logice. Foarte curînd s-a observat că folosirea nelimitată a noțiunii de „mulțime" va duce în mod inevitabil la contradicții. Unul din paradoxuri, indicat de Bertrand Russell, poate fi formulat în felul următor. Cele mai multe mulțimi nu se conțin ca elemente. De exemplu, mulțimea A a tuturor întregilor conține ca elemente numai întregi; A, fiind ea însăși nu un întreg, ci o *mulțime de întregi*, nu se conține ca element. O astfel de mulțime poate fi numită „ordinară". Pot exista mulțimi care se conțin ca elemente. De exemplu, mulțimea S definită în felul următor: „S conține ca elemente toate mulțimile care pot fi definite printr-o frază, care conține mai puțin de treizeci de cuvinte", ar putea fi considerată că se conține ca element. Astfel de mulțimi ar putea fi numite mulțimi „extraordinare". În orice caz, cele mai multe mulțimi sînt ordinare și putem deci exclude comportarea rătăcitoare a mulțimilor „extraordinare", concentrîndu-ne atenția asupra *mulțimii tuturor mulțimilor ordinare*. Fie C această mulțime. Fiecare element al mulțimii C este el însuși o mulțime; și anume o mulțime ordinară. Se pune acum întrebarea, dacă C este ea însăși o mulțime ordinară sau o mulțime extraordinară? Evident, trebuie să fie de un fel sau de altul. Dacă C este ordinară, ea se conține ca element, deoarece C, prin definiție, conține toate mulțimile ordinare. Așa stînd lucrurile, C trebuie să fie extraordinară, deoarece mulțimile extraordinare sînt acelea care se conțin ca element. Aceasta este o contradicție. Prin urmare, C trebuie să fie extraordinară. Dar atunci C conține ca element o mulțime extraordinară (și anume pe ea însăși), ceea ce contrazice definiția prin care C trebuie să conțină numai mulțimi ordinare. Vedem astfel că chiar numai ipoteza existenței mulțimii C ne-a dus la o contradicție.

6. Fundamentele matematicii

Paradoxuri ca acesta l-au dus pe Russell şi pe alţii la un studiu sistematic al fundamentelor matematicii şi logicii. Scopul final al eforturilor lor este de a crea raţionamentului matematic o bază logică lipsită de contradicţii şi care să cuprindă tot ceea ce se consideră a fi important de către toţi (sau unii) matematicieni. În timp ce acest scop ambiţios nu a fost atins şi poate nu va fi atins niciodată, obiectul logicii matematice a atras atenţia unui număr din ce în ce mai mare de oameni de ştiinţă. Multe probleme în acest domeniu, care pot fi enunţate în termeni foarte simpli, sînt foarte greu de rezolvat. Ca exemplu, menţionăm *ipoteza continuului*, care spune că nu există nici o mulţime al cărei număr cardinal este mai mare decît acela al mulţimii întregilor, dar mai mic decît acela al mulţimii numerelor reale. Multe consecinţe interesante pot fi deduse din această ipoteză, dar pînă acum ea nu a fost nici demonstrată, nici infirmată, cu toate că, de curînd, Kurt Gödel a demonstrat că dacă axiomele obişnuite care stau la baza teoriei mulţimilor sînt necontradictorii, atunci sistemul lărgit de axiome, obţinut prin adăugarea ipotezei continuului, este, de asemenea, necontradictoriu[3]. Întrebări ca aceasta se reduc, în cele din urmă, la întrebarea ce se înţelege prin noţiunea de *existenţă* în matematică. Din fericire, existenţa matematicii nu depinde de găsirea unui răspuns satisfăcător la această întrebare. Şcoala „formaliştilor", condusă de marele matematician Hilbert, afirmă că în matematică „existenţa" înseamnă pur şi simplu „absenţa contradicţiei". Atunci devine necesar să se construiască un sistem de axiome, din care să se poată deduce toată matematica, prin raţionament pur formal şi să se arate că acest sistem de axiome nu va duce niciodată la o contradicţie. Rezultate recente ale lui Gödel şi ale altora par să arate că acest program, cel puţin aşa cum a fost conceput la început de Hilbert, nu poate fi efectuat. Este semnificativ faptul că teoria lui Hilbert a structurii formalizate a matematicii se bazează în mod esenţial pe un procedeu intuitiv. Într-un mod sau în altul, pe faţă sau pe ascuns, chiar sub aspectul formalistic, logic sau axiomatic cel mai lipsit de compromis, intuiţia constructivă rămîne întotdeauna elementul vital al matematicii.

§ 5. NUMERELE COMPLEXE

1. Originea numerelor complexe

Din nenumărate motive, noţiunea de număr a trebuit să fie extinsă chiar dincolo de continuul numerelor reale, prin introducerea aşa-numitelor numere complexe. Trebuie să înţelegem că în dezvoltarea istorică şi psihologică a

[3] Problema a fost rezolvată recent de matematicianul american P. Cohen, care a demonstrat că în sistemul de axiome Zermelo—Fraenkel, ipoteza continuului este independentă. — *N.T.*

matematicii, toate aceste extinderi și invenții noi nu erau cîtuși de puțin rezultatul unor eforturi individuale. Ele apar mai degrabă ca rezultat al unei evoluții treptate și ezitante în care nu trebuie exagerat rolul unor persoane individuale. Necesitatea unei mai mari libertăți în calculele formale a fost aceea care a dus la folosirea numerelor negative și raționale. Numai spre sfîrșitul evului mediu matematicienii au început să piardă sentimentul de neliniște și neîncredere cu care operau aceste noțiuni, care nu au părut să aibă același caracter intuitiv concret ca și numerele naturale. Abia spre mijlocul secolului al XIX-lea, matematicienii au înțeles pe deplin că baza logică și filozoifcă esențială peutru operarea într-un domeniu numeric extins este formalistă; că extinderile trebuie să fie create prin definiții care, ca atare, sînt libere, dar care sînt nefolositoare dacă nu sînt făcute în așa fel încît regulile și proprietățile dominante ale sistemului inițial să rămînă în vigoare în sistemul extins. Faptul că aceste extinderi pot fi legate uneori de obiecte „reale" și că în acest fel oferă instrumente pentru noi aplicații este de cea mai mare importanță, dar acesta poate furniza doar o motivare și nu o demonstrație logică a valabilității extinderii.

Procedeul cel mai simplu care necesită folosirea numerelor complexe este acela al *rezolvării ecuațiilor pătratice*. Reamintim noțiunea de ecuație liniară $ax = b$, în care cantitatea necunoscută x trebuie determinată. Soluția este $x = b/a$ și condiția ca orice ecuație liniară cu coeficienți întregi $a \neq 0$ și b să aibă o soluție a impus introducerea numerelor raționale. Ecuații de forma

$$x^2 = 2, \tag{1}$$

care nu au soluții x în cîmpul numerelor raționale, ne-a condus la construirea cîmpului mai larg al numerelor reale, în care există o soluție. Dar chiar cîmpul numerelor reale nu este destul de larg pentru a permite o teorie completă a ecuațiilor pătratice. De exemplu, următoarea ecuație simplă

$$x^2 = -1 \tag{2}$$

nu are soluție reală, deoarece pătratul oricărui număr real nu este niciodată negativ.

Trebuie, fie să fim mulțumiți cu afirmația că această ecuație simplă nu este rezolvabilă, sau să urmăm drumul cunoscut, extinzînd noțiunea noastră de număr prin introducerea numerelor care vor face ecuația rezolvabilă. Tocmai acest lucru se și face cînd introducem noul simbol i, definind $i^2 = -1$. Desigur, acest obiect i — „unitatea imaginară" nu are nimic comun cu noțiunea de număr ca mijloc de *numărare*. El este doar un *simbol* supus regulei fundamentale $i^2 = -1$ și valoarea lui depinde doar de faptul dacă prin această introducere se poate efectua o extindere într-adevăr folositoare și operantă a sistemului de numere.

Deoarece dorim să adunăm și să înmulțim cu simbolul i ca și cu orice număr real, ar trebui să putem forma simboluri ca $2i$, $3i$, $-i$, $2 + 5i$ sau,

mai general, $a + bi$, unde a şi b sînt numere reale oarecare. Dacă aceste simboluri trebuie să asculte de legile obişnuite ale comutativităţii, asociativităţii şi distributivităţii adunării şi înmulţirii, atunci, de exemplu,

$$(2 + 3i) + (1 + 4i) = (2 + 1) + (3 + 4)i = 3 + 7i,$$

$$(2 + 3i)(1 + 4i) = 2 + i + 3i + 12i^2 = (2 - 12) + (8 + 3)i = -10+11i$$

Conduşi de aceste consideraţii, începem expunerea noastră sistematică cu următoarea *definiţie :* un simbol de forma $a + bi$, unde a şi b sînt două numere reale oarecare, se va numi *număr complex* cu *partea reală* a şi *partea imaginară* b. Operaţiile de adunare şi înmulţire vor fi efectuate cu aceste simboluri, ca şi cum i ar fi un număr real obişnuit, cu singura condiţie ca i^2 să fie înlocuit întotdeauna cu -1. Mai precis, definim adunarea şi înmulţirea numerelor complexe prin regulile

$$\begin{aligned} &(a + bi) + (c + di) = (a + c) + (b + d)i, \\ &(a + bi)(c + di) = (ac - bd) + (ad + bc)i. \end{aligned} \tag{3}$$

În particular avem :

$$(a + bi)(a - bi) = a^2 - abi + abi - b^2i^2 = a^2 + b^2. \tag{4}$$

Pe baza acestor definiţii este uşor de verificat că legile comutativităţii, asociativităţii şi distributivităţii sînt valabile pentru numerele complexe. Mai mult, nu numai adunarea şi înmulţirea, dar şi scăderea şi împărţirea a două numere complexe duc din nou la numere de forma $a + bi$, astfel încît numerele complexe formează un *cîmp* (cf. p. 72.):

$$\begin{aligned} &(a + bi) - (c + di) = (a - c) + (b - d)i, \\ &\frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \left(\frac{ac + bd}{c^2 + d^2}\right) + \left(\frac{bc - ad}{c^2 + d^2}\right) i. \end{aligned} \tag{5}$$

(A doua egalitate nu are sens dacă $c + di = 0 + 0i$, deoarece în acest caz $c^2 + d^2 = 0$. Astfel, din nou *trebuie să excludem împărţirea cu zero*, adică cu $0 + 0i$.)

De exemplu :

$$(2 + 3i) - (1 + 4i) = 1 - i,$$

$$\frac{2 + 3i}{1 + 4i} = \frac{2 + 3i}{1 + 4i} \cdot \frac{1 - 4i}{1 - 4i} = \frac{2 - 8i + 3i + 12}{1 + 16} = \frac{14}{17} - \frac{5}{17}\, i.$$

Cîmpul numerelor complexe include cîmpul numerelor reale ca subcîmp, deoarece numărul complex $a + 0i$ este privit ca fiind identic cu numărul

real a. Pe de altă parte, un număr complex de forma $0 + bi = bi$ se numește număr pur imaginar.

Exerciții: 1) Exprimați $\frac{(1+i)(2+i)(3+i)}{(1-i)}$ sub forma $a + bi$.

2) Exprimați

$$\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)^3$$

sub forma $a + bi$.

3) Exprimați sub forma $a + bi$ următoarele numere:

$$\frac{1+i}{1-i},\ \frac{1+i}{2-i},\ \frac{1}{i^6},\ \frac{1}{(-2+i)(1-3i)},\ \frac{(4-5i)^2}{(2-3i)^2}.$$

4) Calculați $\sqrt{5+12i}$. (Indicație: Scrieți $\sqrt{5+12i} = x + yi$, ridicați la pătrat și egalați părțile reale și imaginare.)

Prin introducerea simbolului i am extins cîmpul numerelor reale la un cîmp de simboluri $a + bi$, în care ecuația pătratică particulară

$$x^2 = -1,$$

are soluțiile $x = i$ și $x = -i$. Pentru că, prin definiție, $i \cdot i = (-i)(-i) = i^2 = -1$. În realitate, am cîștigat mai mult: putem verifica acum cu ușurință că *orice ecuație pătratică*, pe care o putem scrie sub forma

$$ax^2 + bx + c = 0, \tag{6}$$

are o soluție. Într-adevăr, din (6) avem

$$x^2 + \frac{b}{a}x = -\frac{c}{a}, \tag{7}$$

$$x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} = \frac{b^2}{4a^2} - \frac{c}{a},$$

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2},$$

$$x + \frac{b}{2a} = \frac{\pm\sqrt{b^2 - 4ac}}{2a}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Acum dacă $b^2 - 4ac \geqslant 0$, atunci $\sqrt{b^2 - 4ac}$ este un număr real obișnuit și soluțiile (7) sînt reale, în timp ce, dacă $b^2 - 4ac < 0$, atunci $4ac - b^2 > 0$ și $\sqrt{b^2 - 4ac} = \sqrt{-(4ac - b^2)} = \sqrt{4ac - b^2} \cdot i$, astfel încît soluțiile (7) sînt numere complexe. De exemplu, soluțiile ecuației

$$x^2 - 5x + 6 = 0$$

sînt $x = (5 \pm \sqrt{25 - 24})/2 = (5 \pm 1)/2 = 2$ sau 3, în timp ce soluțiile ecuației

$$x^2 - 2x + 2 = 0,$$

$$\text{sînt } x = (2 \pm \sqrt{4 - 8})/2 = (2 \pm 2i)/2 = 1 + i \text{ sau } 1 - i.$$

2. Interpretarea geometrică a numerelor complexe

Încă din secolul al XVI-lea matematicienii au fost siliți să introducă expresii pentru rădăcinile pătrate ale numerelor negative, cu scopul de a rezolva toate ecuațiile pătratice și cubice. Ei nu au fost însă în stare să explice înțelesul exact al acestor expresii, pe care ei le priveau cu o înfiorare superstițioasă. Numele „imaginar" ne reamintește de faptul că aceste expresii erau considerate a fi oarecum artificiale, lipsite de valoare reală. În sfîrșit, la începutul secolului al XIX-lea, cînd importanța acestor numere în multe ramuri ale matematicii a devenit evidentă, a fost propusă o interpretare geometrică simplă a operațiilor cu numere complexe, și cu aceasta s-a pus capăt îndoielilor asupra valabilității lor. Desigur, din punct de vedere modern operațiile formale cu numerele complexe se bazează în întregime pe definițiile formale ale adunării și înmulțirii, astfel că interpretarea geometrică nu este logic necesară. Însă interpretarea geometrică, dată aproape concomitent de Wessel (1745—1818), Argand (1768—1822) și Gauss a făcut ca aceste operații să pară mai naturale din punct de vedere intuitiv și de atunci a fost de cea mai mare importanță în aplicațiile numerelor complexe în matematică și fizică.

Această interpretare geometrică constă din reprezentarea numărului complex $z = x + yi$, prin punctul din plan cu coordonatele rectangulare x, y. Astfel, partea reală a lui z este coordonata x, iar partea imaginară este coordonata y. În acest fel s-a stabilit o corespondență între numerele complexe și punctele unui „plan numeric", tot așa cum a fost stabilită în § 2 o corespondență între numerele reale și punctele de pe o dreaptă, axa numerică. Punctele de pe axa Ox a planului numeric corespund numerelor reale $z = x + 0i$, în timp ce punctele de pe axa Oy corespund numerelor pur imaginare $z = 0+yi$.

Dacă

$$z = x + yi$$

este un număr complex, numim numărul complex

$$\bar{z} = x - yi$$

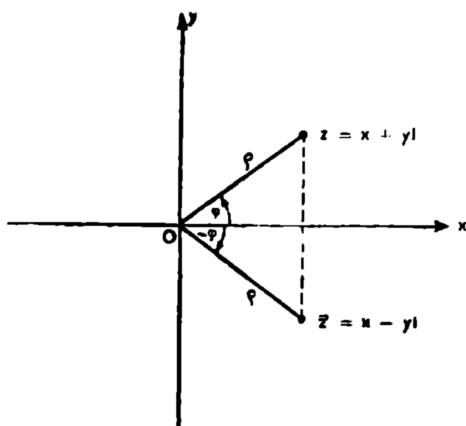


Fig. 22. Reprezentarea geometrică a numerelor complexe. Punctul z are coordonatele rectangulare x, y

conjugatul lui z. Punctul $\bar{z}$ este reprezentat în planul numeric de simetricul punctului z față de axa Ox. Dacă notăm cu ρ distanța de la punctul z la origine, atunci, în virtutea teoremei lui Pitagora

$$\rho^2 = x^2 + y^2 = (x + yi)(x - yi) = z \cdot \bar{z}.$$

Numărul real $\rho = \sqrt{x^2 + y^2}$ se numește *modulul* lui z și se scrie

$$\rho = |z|.$$

Dacă z se află pe axa reală, modulul său este egal cu valoarea absolută obișnuită. Numerele complexe de modul 1 se află pe „cercul unitate" cu centrul în origine și de rază 1.

Dacă $|z| = 0$, atunci $z = 0$. Aceasta rezultă din definiția lui $|z|$ ca distanță de la z la origine. Mai mult, *modulul produsului a două numere complexe este egal cu produsul modulelor lor:*

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|.$$

Aceasta va rezulta dintr-o teoremă mai generală care va fi demonstrată la p. 112.

Exerciții: 1. Demonstrați această teoremă pornind direct de la definiția înmulțirii a două numere complexe $z_1 = x_1 + y_1 i$ și $z_2 = x_2 + y_2 i$.

2. Din faptul că produsul a două numere *reale* este nul numai dacă unul din factori este nul, demonstrați teorema corespunzătoare pentru numere *complexe*. (Indicație: Folosiți cele două teoreme pe care le-am enunțat.)

Din definiția adunării a două numere complexe $z_1 = x_1 + y_1 i$ și $z_2 = x_2 + y_2 i$ avem

$$z_1 + z_2 = (x_1 + x_2) + (y_1 + y_2)i.$$

Deci punctul $z_1 + z_2$ este reprezentat în planul numeric prin cel de-al patrulea vîrf al paralelogramului, ale cărui prime trei vîrfuri sînt punctele O, z_1, z_2. Această construcție geometrică simplă pentru adunarea a două numere

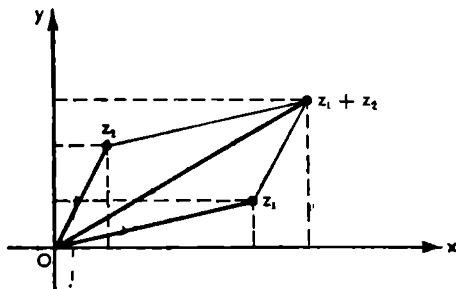


Fig. 23. Regula paralelogramului pentru adunarea numerelor complexe

complexe are mare importanță în multe aplicații. Din ea putem deduce consecința importantă că *modulul sumei a două numere complexe nu întrece suma modulelor*. (Compară cu p. 73.)

$$|z_1 + z_2| \leqslant |z_1| + |z_2|.$$

Aceasta rezultă din faptul că lungimea oricărei laturi a unui triunghi nu poate întrece suma lungimilor celorlalte două laturi.

Exercițiu: Cînd are loc egalitatea $|z_1 + z_2| = |z_1| + |z_2|$?

Unghiul făcut de semiaxa pozitivă Ox cu semidreapta Oz se numește *argumentul* lui z și se notează cu φ (fig. 22). Modulul lui $\bar{z}$ este același cu modulul lui z,

$$|\bar{z}| = |z|,$$

însă argumentul lui $\bar{z}$ este opusul argumentului lui z

$$\bar{\varphi} = -\varphi.$$

Desigur argumentul lui z nu este determinat în mod unic, deoarece i se poate aduna sau scădea orice unghi, multiplu întreg al lui 360°, fără a modifica poziția celei de-a doua laturi. Astfel

$$\varphi, \varphi + 360°, \varphi + 720°, \varphi + 1080°, \ldots$$
$$\varphi - 360°, \varphi - 720°, \varphi - 1080°, \ldots$$

reprezintă toate același argument, din punct de vedere grafic. Cu ajutorul modulului ρ și al argumentului φ, numărul complex z poate fi scris sub forma

$$z = x + yi = \rho(\cos\varphi + i\sin\varphi); \tag{8}$$

deoarece, din definiția sinusului și cosinusului (vezi p. 294),

$$x = \rho\cos\varphi, \qquad y = \rho\sin\varphi.$$

De exemplu, pentru $z = i$, $\rho = 1$, $\varphi = 90°$, astfel încît

$$i = 1\ (\cos 90° + i\sin 90°);$$

pentru $z = 1 + i$, $\rho = \sqrt{2}$, $\varphi = 45°$, astfel încît

$$1 + i = \sqrt{2}(\cos 45° + i\sin 45°);$$

pentru $z = 1 - i$, $\rho = \sqrt{2}$, $\varphi = -45°$, astfel încît

$$1 - i = \sqrt{2}\,[\cos(-45°) + i\sin(-45°)];$$

pentru $z = -1 + \sqrt{3}\,i$, $\rho = 2$, $\varphi = 120°$, astfel încît

$$-1 + \sqrt{3}\,i = 2(\cos 120° + i\sin 120°).$$

Cititorul ar trebui să verifice aceste afirmații substituind valorile funcțiilor trigonometrice.

Reprezentarea trigonometrică (8) are o mare importanță în cazul în care trebuie înmulțite două numere complexe. Dacă

$$z = \rho(\cos\varphi + i\sin\varphi),$$

și

$$z' = \rho'(\cos\varphi' + i\sin\varphi'),$$

atunci

$$zz' = \rho\rho'\{(\cos\varphi\cos\varphi' - \sin\varphi\sin\varphi') + i(\cos\varphi\sin\varphi' + \sin\varphi\cos\varphi')\}.$$

Acum, în baza teoremelor fundamentale de adunare pentru sinus şi cosinus, obţinem:

$$\cos\varphi\cos\varphi' - \sin\varphi\sin\varphi' = \cos(\varphi+\varphi'),$$

$$\cos\varphi\sin\varphi' + \sin\varphi\cos\varphi' = \sin(\varphi+\varphi').$$

Deci

(9) $$zz' = \rho\rho'\{\cos'(\varphi+\varphi') + i\sin(\varphi+\varphi')\}.$$

Aceasta este forma trigonometrică a numărului complex de modul $\rho\rho'$ şi de argument $\varphi+\varphi'$. Cu alte cuvinte, *pentru a înmulţi două numere complexe, înmulţim modulele lor şi le adunăm argumentele* (fig. 24). Astfel, vedem că înmulţirea numerelor complexe are ceva comun cu rotaţia. Pentru a fi mai precişi, să numim un segment orientat, care porneşte din origine şi se termină în punctul z, *vectorul* z. Atunci $\rho = |z|$ va fi lungimea lui. Fie z'

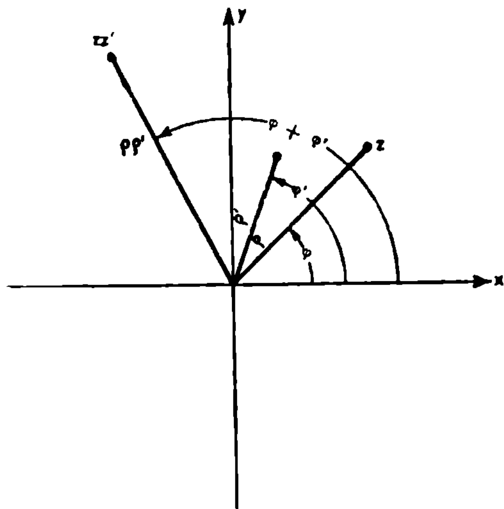


Fig. 24. Înmulţirea a două numere complexe; argumentele se adună, iar modulele se înmulţesc

un număr de pe cercul unitate, astfel încît $\rho' = 1$; atunci, înmulţirea lui z cu z' roteşte pur şi simplu vectorul z cu unghiul φ'. Dacă $\rho' \neq 1$, lungimea vectorului trebuie înmulţită cu ρ' după rotaţie. Cititorul poate ilustra aceste fapte înmulţind diferite numere cu $z_1 = i$ (rotaţie de 90°); $z_2 = -i$ (rotaţie de 90° în sensul opus); $z_3 = 1 + i$; şi $z_4 = 1 - i$.

Formula (9) are o consecinţă deosebit de importantă în cazul în care $z = z'$, deoarece atunci avem

$$z^2 = \rho^2 (\cos 2\varphi + i \sin 2\varphi).$$

Înmulţind acest rezultat din nou cu z, obţinem

$$z^3 = \rho^3(\cos 3\varphi + i \sin 3\varphi),$$

şi continuînd în acest mod găsim

(10) $$z^n = \rho^n (\cos n\varphi + i \sin n\varphi) \quad \text{pentru orice întreg } n.$$

În particular, dacă z este un punct de pe *cercul unitate*, cu $\rho = 1$, obţinem formula descoperită de matematicianul englez A. De Moivre (1667—1754):

(11) $$(\cos \varphi + i \sin \varphi)^n = \cos n\varphi + i \sin n\varphi.$$

Această formulă este una dintre cele mai remarcabile şi folositoare relaţii din matematica elementară. Un exemplu va ilustra acest fapt. Putem aplica formula pentru $n = 3$ şi să dezvoltăm membrul stîng aplicînd formula binomului,

$$(u + v)^3 = u^3 + 3u^2v + 3uv^2 + v^3,$$

obţinînd relaţia

$$\cos 3\varphi + i \sin 3\varphi = \cos^3\varphi - 3 \cos \varphi \sin^2\varphi + i(3 \cos^2\varphi \sin \varphi - \sin^3\varphi).$$

O singură egalitate, ca aceasta, *între două numere complexe*, *este echivalentă cu o pereche de egalităţi între numere reale.* Pentru că, dacă două numere complexe sînt egale, atît părţile reale cît şi cele imaginare trebuie să fie egale. Deci putem scrie

$$\cos 3\varphi = \cos^3\varphi - 3 \cos \varphi \sin^2\varphi, \qquad \sin 3\varphi = 3 \cos^2\varphi \sin \varphi - \sin^3\varphi.$$

Folosind relaţia

$$\cos^2\varphi + \sin^2\varphi = 1,$$

avem în cele din urmă

$$\cos 3\varphi = \cos^3\varphi - 3 \cos\varphi(1 - \cos^2\varphi) = 4 \cos^3\varphi - 3 \cos \varphi,$$

$$\sin 3\varphi = -4 \sin^3\varphi + 3 \sin \varphi.$$

Formule asemănătoare, care exprimă pe $\sin n\varphi$ şi $\cos n\varphi$ în funcţie de puterile lui $\sin \varphi$ şi $\cos \varphi$, pot fi obţinute cu uşurinţă pentru orice valoare a lui n.

Exerciții: 1) Găsiți formulele corespunzătoare pentru sin 4φ și cos 4φ.

2) Demonstrați că pentru un punct $z = \cos\varphi + i\sin\varphi$, de pe cercul unitate, avem $1/z = \cos\varphi - i\sin\varphi$.

3) Demonstrați, fără a calcula, că $(a + bi)/(a - bi)$ are întotdeauna valoarea absolută egală cu 1.

4) Dacă z_1 și z_2 sînt două numere complexe, arătați că argumentul lui $z_1 - z_2$ este egal cu unghiul făcut de axa reală cu vectorul care pornește din z_2 și se termină în z_1.

5) Interpretați argumentul numărului complex $(z_1 - z_2)/(z_1 - z_3)$ în triunghiul format de punctele z_1, z_2 și z_3.

6) Demonstrați că cîtul a două numere complexe care au același argument este real.

7) Demonstrați că dacă pentru patru numere complexe z_1, z_2, z_3, z_4 argumentele lui $\frac{z_3 - z_1}{z_3 - z_2}$ și $\frac{z_4 - z_1}{z_4 - z_2}$ sînt egale, atunci cele patru numere se află pe un cerc, sau pe o dreaptă și reciproc.

8) Demonstrați că patru puncte z_1, z_2, z_3, z_4 se află pe un cerc sau pe o dreaptă, dacă și numai dacă raportul

$$\frac{z_3 - z_1}{z_3 - z_2} \Big/ \frac{z_4 - z_1}{z_4 - z_2}$$

este real.

3. Formula lui De Moivre și rădăcinile unității

Prin rădăcină de ordinul n a unui număr a, înțelegem un număr b, astfel încît $b^n = a$. În particular, numărul 1 are rădăcinile pătrate 1 și -1, deoarece $1^2 = (-1)^2 = 1$. Numărul 1 are o singură rădăcină cubică reală, 1, în timp ce el are patru rădăcini de ordinul patru: numerele reale 1 și -1 și numerele imaginare i și $-i$. Aceste fapte sugerează că în domeniul complex s-ar putea să existe încă două rădăcini cubice ale lui 1, deci în total vor fi trei rădăcini cubice. Că lucrurile stau astfel, se poate vedea imediat din formula lui De Moivre.

Vom vedea că *în cîmpul numerelor complexe există exact* n *rădăcini diferite, de ordinul* n, *ale lui* 1. *Ele sînt reprezentate de vîrfurile unui poligon regulat cu* n *laturi, înscris în cercul unitate, unul din vîrfurile lui fiind punctul* z = 1. Aceasta se vede aproape imediat din fig. 25 (desenată pentru cazul $n = 12$). Primul vîrf al poligonului este 1. Următorul este

$$\alpha = \cos\frac{360°}{n} + i\sin\frac{360°}{n}, \tag{12}$$

deoarece argumentul său trebuie să fie egal cu a n-a parte a unghiului total de 360°. Următorul vîrf este $\alpha \cdot \alpha = \alpha^2$, deoarece îl obținem prin rotirea

vectorului α cu unghiul $\frac{360°}{n}$. Următorul vîrf este α^3 etc., şi, în sfîrşit, după n paşi sîntem din nou în vîrful 1, adică avem

$$\alpha^n = 1,$$

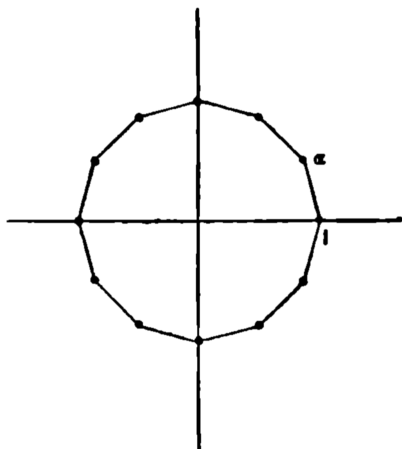


Fig. 25. Cele douăsprezece rădăcini de ordinul doisprezece ale lui 1.

ceea ce rezultă de asemenea din formula (11), deoarece

$$\left[\cos\frac{360°}{n} + i\ \sin\frac{360°}{n}\right]^n = \cos 360° + i \sin 360° = 1 + 0i.$$

Rezultă că $\alpha^1 = \alpha$ este o rădăcină a ecuaţiei $x^n = 1$. Acelaşi lucru este adevărat pentru vîrful următor $\alpha^2 = \cos\left(\frac{720°}{n}\right) + i\ \sin\left(\frac{720°}{n}\right)$. Putem vedea acest lucru scriind

$$(\alpha^2)^n = \alpha^{2n} = (\alpha^n)^2 = (1)^2 = 1,$$

sau, din formula lui De Moivre:

$$(\alpha^2)^n = \cos\left(n\,\frac{720°}{n}\right) + i\ \sin\left(n\,\frac{720°}{n}\right) = \cos 720° + i \sin 720° =$$

$$= 1 + 0i = 1.$$

În acelaşi mod, vedem că toate cele n numere

$$1, \alpha, \alpha^2, \alpha^3, \ldots, \alpha^{n-1}$$

sînt rădăcini de ordinul n ale lui 1. Mergînd mai departe în şirul exponenţilor sau folosind exponenţi negativi, nu obţinem rădăcini noi. Într-adevăr $\alpha^{-1} = 1/\alpha = \alpha^n/\alpha = \alpha^{n-1}$ şi $\alpha^n = 1$, $\alpha^{n+1} = (\alpha)^n\alpha = 1 \cdot \alpha = \alpha$ etc., astfel încît valorile precedente se repetă. Lăsăm în sarcina cititorului să arate că nu există alte rădăcini de ordinul n.

Dacă n este par, atunci unul din vîrfurile poligonului cu n laturi se va afla în punctul -1, în concordanţă cu faptul algebric că în acest caz -1 este rădăcină de ordinul n a lui 1.

Ecuaţia satisfăcută de rădăcinile de ordinul n ale lui 1

$$x^n - 1 = 0 \tag{13}$$

este de gradul n, dar ea poate fi redusă cu uşurinţă la o ecuaţie de gradul $(n - 1)$, folosind formula algebrică

$$x^n - 1 = (x - 1)(x^{n-1} + x^{n-2} + x^{n-3} + \ldots + 1). \tag{14}$$

Deoarece produsul a două numere este nul dacă *şi numai dacă* cel puţin unul din cele două numere este nul, membrul stîng al egalităţii (14) se anulează, numai dacă unul din cei doi factori din membrul drept este nul, adică numai dacă $x = 1$, sau dacă este satisfăcută ecuaţia

$$x^{n-1} + x^{n-2} + x^{n-3} + \ldots + x + 1 = 0. \tag{15}$$

Rezultă deci că aceasta este ecuaţia care trebuie să fie satisfăcută de rădăcinile $\alpha, \alpha^2, \ldots, \alpha^{n-1}$; ea se numeşte *ecuaţia ciclotomică* (a diviziunii cercului). De exemplu, rădăcinile cubice complexe ale lui 1,

$$\alpha = \cos 120° + i \sin 120° = \frac{1}{2}(-1 + i\sqrt{3}),$$

$$\alpha^2 = \cos 240° + i \sin 240° = \frac{1}{2}(-1 - i\sqrt{3}),$$

sînt rădăcinile ecuaţiei

$$x^2 + x + 1 = 0,$$

după cum cititorul poate vedea imediat prin substituţie directă. În mod analog, rădăcinile de ordinul 5 ale lui 1, diferite de 1, satisfac ecuaţia

$$x^4 + x^3 + x^2 + x + 1 = 0. \tag{16}$$

Pentru a construi un pentagon regulat trebuie să rezolvăm această ecuație de gradul patru. Printr-un artificiu algebric simplu ea poate fi redusă la o ecuație pătratică în raport cu $w = x + 1/x$. Împărțim ecuația (16) prin x^2 și reordonăm termenii:

$$x^2 + \frac{1}{x^2} + x + \frac{1}{x} + 1 = 0,$$

sau, deoarece $(x + 1/x)^2 = x^2 + 1/x^2 + 2$, obținem ecuația

$$w^2 + w - 1 = 0.$$

Din formula (7) a secțiunii 1 rezultă că această ecuație are rădăcinile

$$w_1 = \frac{-1 + \sqrt{5}}{2}, \quad w_2 = \frac{-1 - \sqrt{5}}{2}.$$

Deci, rădăcinile complexe de ordinul 5 ale lui 1 sînt rădăcinile a două ecuații de ordinul doi

$$x + \frac{1}{x} = w_1; \text{ sau } x^2 + \frac{1}{2}(\sqrt{5} - 1)x + 1 = 0$$

și

$$x + \frac{1}{x} = w_2, \text{ sau } x^2 - \frac{1}{2}(\sqrt{5} + 1)x + 1 = 0,$$

pe care cititorul le poate rezolva cu ajutorul formulei pe care tocmai am folosit-o.

Exerciții: 1) Găsiți rădăcinile de ordinul 6 ale lui 1. 2) Găsiți $(1 + i)^{11}$. 3) Găsiți toate valorile lui $\sqrt{1 + i}$, $\sqrt[3]{7 - 4i}$, $\sqrt[3]{i}$, $\sqrt[5]{-i}$. 4) Calculați $(1/2i)(i^7 - i^{-7})$.

*4. Teorema fundamentală a algebrei

Nu numai orice ecuație de forma $ax^2 + bx + c = 0$, sau de forma $x^n - 1 = 0$, este rezolvabilă în cîmpul numerelor complexe, dar mai mult: *orice ecuație algebrică de orice grad* n, *cu coeficienți reali sau complecși,*

(17) $$f(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \ldots + a_1x + a_0 = 0,$$

are soluții în cîmpul numerelor complexe. Pentru ecuația de gradul 3 și 4, acest lucru a fost stabilit în secolul al XVI-lea de Tartaglia, Cardan și alții, care au rezolvat astfel de ecuații cu ajutorul unor formule asemănătoare în esență cu aceea pentru rezolvarea ecuației pătratice, cu toate că erau

mult mai complicate. Timp de aproape două sute de ani, ecuațiile generale de gradul 5 și de grade superioare au fost studiate intens, dar toate eforturile făcute pentru a le rezolva prin metode asemănătoare au dat greș. A fost o mare realizare atunci cînd tînărul Gauss, în teza sa de doctorat (1799), a reușit să dea prima demonstrație completă a faptului că soluțiile *există*, cu toate că problema generalizării formulelor clasice, care exprimă soluțiile ecuațiilor de grad mai mic decît 5, cu ajutorul unor operații raționale și al extragerii rădăcinii, rămînea nerezolvată chiar și atunci (cf. p. 135).

Teorema lui Gauss afirmă că *pentru orice ecuație algebrică de forma* (17), *unde* n *este un întreg pozitiv iar coeficienții* a *sînt numere reale sau complexe oarecare, există cel puțin un număr complex* $\alpha = c + di$, *astfel încît*

$$f(\alpha) = 0.$$

Numărul α se numește *rădăcină* a ecuației (17). O demonstrație a acestei teoreme va fi dată la p. 286. Admițînd pentru moment valabilitatea ei, putem demonstra așa-numita *teoremă fundamentală a algebrei* (ar trebui să fie numită mai potrivit teorema fundamentală a sistemului de numere complexe). *Orice polinom de gradul n*

$$(18) \qquad f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0,$$

poate fi descompus într-un produs de n *factori*,

$$(19) \qquad f(x) = (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_n),$$

unde $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$ *sînt numere complexe, rădăcinile ecuației* $f(x) = 0$. Astfel, de exemplu, polinomul

$$f(x) = x^4 - 1$$

poate fi descompus sub forma

$$f(x) = (x - 1)(x - i)(x + i)(x + 1).$$

Faptul că numerele α sînt rădăcinile ecuației $f(x) = 0$ este evident din descompunerea (19), deoarece pentru $x = \alpha_r$, unul din factorii lui $f(x)$, și deci și $f(x)$, este egal cu zero.

În unele cazuri, factorii $(x - \alpha_1)$, $(x - \alpha_2)$, ... ai unui polinom $f(x)$ de gradul n nu sînt toți distincți, ca în exemplul

$$f(x) = x^2 - 2x + 1 = (x - 1)(x - 1),$$

care are o singură rădăcină $x = 1$, „numărată de două ori", sau „de multiplicitate 2". În orice caz, un polinom de gradul n nu poate avea mai mult de n factori distincți $(x - \alpha)$ și ecuația corespunzătoare nu poate avea mai mult de n rădăcini.

Pentru a demonstra teorema descompunerii în factori folosim din nou identitatea algebrică

$$(20)\quad x^k - \alpha^k = (x - \alpha)(x^{k-1} + \alpha x^{k-2} + \alpha^2 x^{k-3} + \ldots + \alpha^{k-2}x + \alpha^{k-1}),$$

care pentru $\alpha = 1$ se reduce la formula pentru progresia geometrică. Deoarece admitem valabilitatea teoremei lui Gauss, putem presupune că $\alpha = \alpha_1$ este o rădăcină a ecuației (17), astfel încît

$$f(\alpha_1) = \alpha_1^n + a_{n-1}\alpha_1^{n-1} + a_{n-2}\alpha_1^{n-2} + \ldots + a_1\alpha_1 + a_0 = 0.$$

Scăzînd aceasta din $f(x)$ și reordonînd termenii, obținem identitatea

$$(21)\quad f(x) = f(x) - f(\alpha_1) = (x^n - \alpha_1^n) + a_{n-1}(x^{n-1} - \alpha_1^{n-1}) + \ldots$$
$$\ldots + a_1(x - \alpha_1).$$

Acum, datorită identității (20) putem extrage factorul $(x - \alpha_1)$ din fiecare termen al egalității (21), astfel încît gradul celuilalt factor al fiecărui termen este micșorat cu 1. Deci, reordonînd din nou termenii, găsim că

$$f(x) = (x - \alpha_1)g(x),$$

unde $g(x)$ este un polinom de gradul $n - 1$:

$$g(x) = x^{n-1} + b_{n-2}x^{n-2} + \ldots + b_1 x + b_0.$$

(Pentru scopurile noastre nu este necesar să calculăm coeficienții b_k.) Mai departe aplicăm același procedeu lui $g(x)$. Din teorema lui Gauss rezultă că există o rădăcină α_2 a ecuației $g(x) = 0$, astfel încît

$$g(x) = (x - \alpha_2)h(x),$$

unde $h(x)$ este un polinom de gradul $n - 2$. Continuînd în acest mod de $(n - 1)$ ori (desigur, această frază substituie un raționament prin inducție matematică), obținem în cele din urmă descompunerea completă

$$22)\quad f(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \ldots (x - \alpha_n).$$

Din (22) rezultă nu numai că numerele complexe $\alpha_1, \alpha_2, \ldots, \alpha_n$ sînt rădăcinile ecuației (17), dar că ele sînt *singurele* rădăcini. Într-adevăr, dacă y ar fi o rădăcină a ecuației (17), atunci din (22)

$$f(y) = (y - \alpha_1)(y - \alpha_2) \ldots (y - \alpha_n) = 0.$$

Am văzut la p. 111 că un produs de numere complexe este egal cu 0 dacă *și numai dacă* unul din factori este egal cu 0. Deci unul din factorii $(y - \alpha_r)$ trebuie să fie nul, iar y trebuie să fie egal cu α_r, ceea ce trebuia arătat.

*§ 6. NUMERE ALGEBRICE ŞI TRANSCENDENTE

1. Definiţie şi existenţă

Un *număr algebric* este orice număr x, real sau complex, care satisface o ecuaţie algebrică de forma

$$(1) \qquad a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 = 0 \qquad (n \geqslant 1,\ a_n \neq 0),$$

unde a_k sînt *întregi*. De exemplu, $\sqrt{2}$ este un număr algebric, deoarece el satisface ecuaţia

$$x^2 - 2 = 0.$$

În mod analog, orice rădăcină a unei ecuaţii cu coeficienţi întregi, de gradul trei, patru, cinci sau mai mare este un număr algebric, chiar dacă rădăcinile pot fi sau nu exprimate cu ajutorul radicalilor. Noţiunea de număr algebric este o generalizare naturală a numărului raţional, care este un caz particular corespunzător lui $n = 1$.

Nu orice număr real este algebric. Acest lucru se poate vedea cu ajutorul unei demonstraţii datorită lui Cantor, care a arătat că mulţimea numerelor algebrice este *numărabilă*. Deoarece mulţimea tuturor numerelor reale este nenumărabilă, trebuie să existe numere reale care nu sînt algebrice.

O metodă de numărare a mulţimii numerelor algebrice este următoarea: Oricărei ecuaţii de forma (1) i se asociază întregul pozitiv

$$h = |a_n| + |a_{n-1}| + \ldots + |a_1| + |a_0| + n$$

numit „înălţime" a ecuaţiei. Pentru orice valoare *fixată a* lui h există doar un număr *finit* de ecuaţii (1) cu înălţimea h. Fiecare din aceste ecuaţii poate avea cel mult n rădăcini diferite. De aceea există doar un număr finit de numere algebrice ale căror ecuaţii au înălţimea h şi putem aşeza toate numerele algebrice într-un şir începînd cu acelea de înălţime 1, luînd apoi pe acelea de înălţime 2 şi aşa mai departe.

Această demonstraţie a faptului că mulţimea numerelor algebrice este numărabilă asigură existenţa numerelor reale, care nu sînt algebrice; astfel de numere se numesc *transcendente*, deoarece, după cum a spus Euler, ele „depăşesc puterile metodelor algebrice".

Demonstraţia lui Cantor a existenţei numerelor transcendente nu poate fi considerată constructivă. Din punct de vedere teoretic am putea construi un număr transcendent aplicînd procedeul diagonal al lui Cantor unui tabel imaginar de reprezentări zecimale ale tuturor rădăcinilor ecuaţiilor algebrice, dar acest procedeu ar fi cu totul nepractic şi nu ar duce la nici un număr a cărui reprezentare în sistemul zecimal, sau în oricare alt sistem, ar putea

fi scrisă efectiv. Mai mult, cele mai interesante probleme referitoare la numerele transcendente sînt legate de demonstrarea faptului că anumite numere concrete, ca de pildă π și e (aceste numere vor fi definite la pp. 316—317), sînt într-adevăr transcendente.

**2. Teorema lui Liouville și construcția numerelor transcendente

O demonstrație a existenței numerelor transcendente, care precede pe cea a lui Cantor, a fost dată de J. Liouville (1809—1882). Demonstrația lui Liouville permite *construirea* efectivă a unor exemple de astfel de numere. E puțin mai dificilă decît demonstrația lui Cantor, tot așa cum sînt cele mai multe construcții în comparație cu demonstrațiile de existență. Demonstrația este inserată aici numai pentru cititorul mai avansat, deși ea nu necesită decît cunoașterea matematicii elementare.

Liouville a arătat că numerele algebrice iraționale au proprietatea că nu pot fi aproximate prin numere raționale, cu o precizie foarte mare, decît dacă numitorii fracțiilor de aproximare sînt foarte mari.

Să presupunem că numărul z satisface ecuația algebrică cu coeficienți întregi

$$(2) \qquad f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n = 0 \qquad (a_n \neq 0),$$

dar nu satisface o ecuație similară de grad mai mic. Atunci se spune că z este un număr algebric *de gradul n*. De exemplu, $z = \sqrt{2}$ este un număr algebric de gradul doi, deoarece el satisface ecuația $x^2 - 2 = 0$, dar nu satisface nici o ecuație de gradul întîi; $z = \sqrt[3]{2}$ este de gradul trei, pentru că satisface ecuația $x^3 - 2 = 0$ și, așa cum vom vedea în capitolul III, nu satisface nici o ecuație de grad mai mic. Un număr algebric de grad $n > 1$ nu poate fi rațional, deoarece un număr rațional $z = p/q$ satisface ecuația $qx - p = 0$, de gradul întîi. Deci, orice număr irațional z poate fi aproximat, cu orice grad de precizie, printr-un număr rațional; aceasta înseamnă că putem găsi un șir

$$\frac{p_1}{q_1}, \frac{p_2}{q_2}, \ldots$$

de numere raționale, cu numitori din ce în ce mai mari, astfel încît

$$\frac{p_r}{q_r} \to z.$$

Teorema lui Liouville afirmă că pentru orice număr algebric z de grad $n > 1$, o astfel de aproximație are o precizie mai mică decît $1/q^{n+1}$; adică, trebuie să aibă loc inegalitatea

$$(3) \qquad \left| z - \frac{p}{q} \right| > \frac{1}{q^{n+1}}$$

pentru numitori q destul de mari.

Vom demonstra această teoremă, dar mai întîi vom arăta cum permite ea construirea numerelor transcendente. Să luăm numărul (cf. p. 34 pentru definirea simbolului $n!$)

$$z = a_1 \cdot 10^{-1!} + a_2 \cdot 10^{-2!} + a_3 \cdot 10^{-3!} + \ldots + a_m \cdot 10^{-m!} + \\ + a_{m+1} \cdot 10^{-(m+1)!} + \ldots \\ = 0,\, a_1\, a_2\, 000\, a_3\, 00000000000000000\, a_4\, 0000000\ldots,$$

unde coeficienții a_i sînt cifre arbitrare cuprinse între 1 și 9 (de exemplu, am putea lua toți a_i egali cu 1). Un astfel de număr este caracterizat prin succesiuni de 0-uri în creștere rapidă, întrerupte de cîte o singură cifră nenulă. Să notăm cu z_m fracția zecimală finită formată cînd în reprezentare luăm doar termenii lui z pînă la $a_m \cdot 10^{-m!}$, inclusiv. Atunci

$$(4) \qquad |z - z_m| < 10 \cdot 10^{-(m+1)!}.$$

Să presupunem că z ar fi un număr algebric de gradul n. Atunci în (3) să punem $p/q = z_m = p/10^{m!}$; obținem

$$|z - z_m| > \frac{1}{10^{(n+1)m!}}$$

pentru m suficient de mare. Ținînd seama de (4), ar trebui să avem

$$\frac{1}{10^{(n+1)m!}} < \frac{10}{10^{(m+1)!}} = \frac{1}{10^{(m+1)!-1}},$$

astfel încît $(n + 1)m! > (m + 1)! - 1$ pentru orice m suficient de mare. Dar această inegalitate nu este adevărată dacă m depășește pe n (cititorul ar trebui să dea o demonstrație detaliată a acestei afirmații), ceea ce ne dă o contradicție. Deci z este transcendent.

Rămîne de demonstrat teorema lui Liouville. Să presupunem că z este număr algebric de grad $n > 1$, care satisface ecuația (1), adică

$$f(z) = 0. \tag{5}$$

Fie $z_m = p_m/q_m$ un șir de numere raționale pentru care $z_m \to z$. Atunci

$$f(z_m) = f(z_m) - f(z) = a_1(z_m - z) + a_2(z_m^2 - z^2) + \ldots + a_n(z_m^n - z^n).$$

Împărțind ambii membri ai acestei egalități cu $z_m - z$, și folosind formula algebrică

$$\frac{u^n - v^n}{u - v} = u^{n-1} + u^{n-2}v + u^{n-3}v^2 + \ldots + uv^{n-2} + v^{n-1},$$

obținem

$$\frac{f(z_m)}{z_m - z} = a_1 + a_2(z_m + z) + a_3(z_m^2 + z_m z + z^2) + \ldots + \tag{6}$$
$$+ a_n(z_m^{n-1} + \ldots + z^{n-1}).$$

Deoarece z_m are limita z, el va diferi de z cu mai puțin de 1, dacă m este suficient de mare. De aceea, putem scrie următoarea evaluare, pentru m suficient de mare:

$$\left|\frac{f(z_m)}{z_m - z}\right| < |a_1| + 2|a_2|(|z| + 1) + 3|a_3|(|z| + 1)^2 + \ldots \tag{7}$$
$$+ n|a_n|(|z| + 1)^{n-1} = M,$$

care este un număr constant, deoarece și z este constant în raționamentul nostru. Dacă alegem acum pe m atît de mare încît în $z_m = \frac{p_m}{q_m}$, numitorul q_m să fie mai mare decît M, atunci

$$|z - z_m| > \frac{|f(z_m)|}{M} > \frac{|f(z_m)|}{q_m}. \tag{8}$$

Pentru a simplifica notația, să înlocuim pe p_m prin p și pe q_m prin q. Atunci

$$|f(z_m)| = \left|\frac{a_0 q^n + a_1 q^{n-1} p + \ldots + a_n p^n}{q^n}\right|. \tag{9}$$

Numărul rațional $z_m = p/q$ nu poate fi o rădăcină a ecuației $f(x) = 0$, deoarece, dacă ar fi rădăcină, am putea extrage factorul $(x - z_m)$ din $f(x)$, și z ar satisface o ecuație de grad mai mic decît n. Deci $f(z_m) \neq 0$. Însă numărătorul membrului drept din (9) este un întreg și deci el trebuie să fie cel puțin egal cu 1. Deci din (8) și (9) avem:

$$|z - z_m| > \frac{1}{q}\,\frac{1}{q^n} = \frac{1}{q^{n+1}}, \tag{10}$$

ceea ce demonstrează teorema.

În cursul ultimelor decenii, cercetările asupra posibilității aproximării numerelor algebrice prin numere raționale au fost duse mult mai departe. De exemplu, matematicianul norvegian A. Thue (1863—1922) a demonstrat că în inegalitatea (3) a lui Liouville, exponentul $n + 1$ poate fi înlocuit cu $(n/2) + 1$. C. L. Siegel a arătat mai tîrziu că are loc chiar propoziția mai tare (mai tare pentru valori mari ale lui n) cu exponentul $2\sqrt{n}$.

Subiectul numerelor transcendente a fascinat întotdeauna pe matematicieni. Dar pînă de curînd erau cunoscute foarte puține exemple de numere interesante în sine, despre care s-a putut arăta că sînt transcendente. (În capitolul III vom discuta caracterul transcendent al lui π, din care rezultă imposibilitatea cvadraturii cercului cu ajutorul riglei și compasului.) Într-o comunicare celebră la Congresul internațional al matematicienilor, care a avut loc la Paris în 1900, David Hilbert a propus treizeci de probleme matematice care erau ușor de formulat, unele chiar în limbaj elementar și popular, dar dintre care nici una nu era rezolvată sau nu părea a fi imediat accesibilă tehnicii matematice existente pe atunci. Aceste „probleme ale lui Hilbert" au constituit o provocare în perioada de dezvoltare matematică care a urmat. Aproape toate au fost rezolvate între timp și soluția a însemnat adesea un progres cert al concepției matematice și al metodelor generale. Una dintre probleme, care părea să aibă cele mai puține speranțe de rezolvare era să se demonstreze că

$$2^{\sqrt{2}}$$

este un număr transcendent sau chiar irațional. Timp de aproape trei decenii nu a existat nici cea mai mică sugestie pentru o direcție de atac promițătoare pentru această problemă. În sfîrșit, Siegel și, în mod independent, tînărul rus A. Ghelfond au descoperit noi metode de demonstrare a caracterului transcendent al multor numere semnificative din matematică, printre care se află și numărul lui Hilbert $2^{\sqrt{2}}$ și, mai general, orice număr a^b unde a este un număr algebric $\neq 0$ sau 1, iar b este orice număr algebric irațional.

ALGEBRA MULŢIMILOR

1. Teoria generală

Noţiunea de *clasă* sau *mulţime* de obiecte este una dintre noţiunile fundamentale ale matematicii. O mulţime este definită de orice proprietate, sau atribut, $\mathfrak{A}$ pe care fiecare din obiectele considerate trebuie, fie să o posede fie să nu o posede; obiectele care posedă proprietatea formează o mulţime corespunzătoare A. Astfel, dacă considerăm întregii şi dacă proprietatea $\mathfrak{A}$ este de a fi un număr prim, mulţimea corespunzătoare A este mulţimea tuturor numerelor prime 2, 3, 5, 7, ...

Studiul matematic al mulţimilor este bazat pe faptul că mulţimile pot fi combinate prin anumite operaţii, pentru a forma alte mulţimi, tot aşa cum numerele pot fi combinate prin adunare şi înmulţire pentru a forma alte numere. Studiul operaţiilor cu mulţimi cuprinde „algebra mulţimilor", care are multe asemănări formale, ca şi deosebiri, faţă de algebra numerelor. Faptul că se pot aplica metode algebrice studiului unor obiecte nenumerice, cum sînt mulţimile, ilustrează marea generalitate a noţiunilor matematicii moderne. În ultimii ani, a devenit evident faptul că algebra mulţimilor clarifică multe ramuri ale matematicii, ca de pildă teoria măsurii şi teoria probabilităţilor; ea este utilă, de asemenea, pentru reducerea noţiunilor matematice la fundamentul lor logic.

În cele ce urmează, prin I se va desemna o mulţime fixată de obiecte de natură oarecare, numită mulţimea universală, sau universul discursului, iar A, B, C, ... vor desemna submulţimi arbitrare ale lui I. Dacă I desemnează mulţimea tuturor întregilor, A poate desemna mulţimea tuturor întregilor pari, B mulţimea tuturor întregilor impari, C mulţimea tuturor numerelor prime etc. Sau I ar putea să desemneze mulţimea tuturor punctelor unui anumit plan, A mulţimea tuturor punctelor din interiorul unui cerc cuprins în plan, B mulţimea tuturor punctelor din interiorul altui cerc din plan etc. Pentru comoditate, includem printre „submulţimile" lui I mulţimea I şi „mulţimea vidă" $\emptyset$, care nu conţine nici un element. Scopul

acestei extinderi artificiale este de a menține regula ca oricărei proprietăți $\mathfrak{A}$ să-i corespundă submulțimea A, a tuturor elementelor lui I care au această proprietate. În cazul în care $\mathfrak{A}$ este o proprietate universal valabilă, ca de pildă aceea specificată de egalitatea banală $x = x$, submulțimea corespunzătoare a lui I va fi chiar I, deoarece orice obiect satisface această egalitate, în timp ce dacă $\mathfrak{A}$ este o proprietate care se contrazice pe sine, ca de pildă $x \neq x$, submulțimea corespunzătoare nu va conține obiecte și poate fi notată prin simbolul $\emptyset$.

Se spune că mulțimea A este o *submulțime* a mulțimii B dacă nu există nici un obiect în A care să nu se afle în B. Dacă se întîmplă acest lucru, scriem

$$A \subset B \quad \text{sau} \quad B \supset A.$$

De exemplu, mulțimea A a tuturor întregilor care sînt multipli de 10 este o submulțime a mulțimii B a tuturor întregilor care sînt multipli de 5, deoarece orice multiplu de 10 este de asemenea un multiplu de 5. Afirmația $A \subset B$ nu exclude posibilitatea ca $B \subset A$. Dacă ambele relații au loc, spunem că mulțimile A și B sînt egale și scriem

$$A = B.$$

Pentru ca această egalitate să fie adevărată, trebuie ca orice element a lui A să fie un element a lui B, și reciproc, astfel încît mulțimile A și B conțin aceleași elemente.

Relația $A \subset B$ are multe asemănări cu relația de ordine $a \leqslant b$ dintre numerele reale. În particular, este adevărat că

1) $A \subset A$.
2) Dacă $A \subset B$ și $B \subset A$, atunci $A = B$.
3) Dacă $A \subset B$ și $B \subset C$, atunci $A \subset C$.

Pentru acest motiv numim relația $A \subset B$ tot o „relație de ordine". Deosebirea principală față de relația $a \leqslant b$ dintre numere este faptul că în timp ce pentru *orice* pereche de numere a și b cel puțin una din relațiile $a \leqslant b$ sau $b \leqslant a$ are loc întotdeauna, acest lucru nu mai este adevărat pentru mulțimi. De exemplu, dacă A desemnează mulțimea formată din întregii 1, 2, 3

$$A = \{1, 2, 3\},$$

iar B mulțimea formată din întregii 2, 3, 4

$$B = \{2, 3, 4\},$$

atunci nu avem nici $A \subset B$, nici $B \subset A$. Pentru acest motiv se spune că relația $A \subset B$ determină o ordine parțială între mulțimi, în timp ce relația $a \leqslant b$ determină o ordine totală între numere.

În treacăt, putem observa că din definiția relației $A \subset B$ rezultă că

4) $\varnothing \subset A$ pentru orice mulțime A, și,

5) $A \subset I$,

unde A este orice submulțime a universului I. Relația (4) poate părea oarecum paradoxală, dar ea este în concordanță cu interpretarea strictă a definiției semnului $\subset$. Într-adevăr, relația $\varnothing \subset A$ ar putea fi falsă numai dacă mulțimea vidă $\varnothing$ ar conține un obiect care nu se află în A, și deoarece mulțimea vidă nu conține obiecte, acest lucru este imposibil, oricare ar fi mulțimea A.

Vom defini acum două operații cu mulțimi, care formal au multe din proprietățile algebrice ale adunării și înmulțirii obișnuite cu numere, cu toate că din punct de vedere conceptual ele sînt cu totul deosebite de aceste operații. În acest scop, fie A și B două mulțimi oarecare. Prin „reuniune" sau „sumă logică" a lui A și B înțelegem mulțimea care constă din toate obiectele care se află *fie* în A, *fie* în B (inclusiv în ambele). Această mulțime va fi desemnată prin simbolul $A + B$. Prin „intersecția" sau „produsul logic" al lui A și B înțelegem mulțimea care constă numai din acele elemente care se află *atît* în A, *cît* și în B. Desemnăm această mulțime prin simbolul $A \cdot B$ sau mai simplu AB. Pentru a ilustra aceste operații, putem alege din nou ca mulțimi A și B mulțimile

$$A = \{1, 2, 3\}, \qquad B = \{2, 3, 4\}.$$

Atunci

$$A + B = \{1, 2, 3, 4\}, \; AB = \{2, 3\}.$$

Printre proprietățile algebrice importante ale operațiilor $A + B$ și AB enumărăm următoarele. Ele ar trebui să fie verificate de cititor, pe baza definiției acestor operații:

6) $A + B = B + A$	7) $AB = BA$
8) $A + (B + C) = (A + B) + C$	9) $A(BC) = (AB)C$
10) $A + A = A$	11) $AA = A$
12) $A(B + C) = AB + AC$	13) $A + (BC) = (A + B)(A + C)$
14) $A + \varnothing = A$	15) $AI = A$
16) $A + I = I$	17) $A\varnothing = \varnothing$

18) Relația $A \subset B$ este echivalentă cu fiecare din relațiile

$$A + B = B, \; AB = A.$$

Verificarea acestor legi este o chestiune de logică elementară. De exemplu, 10) afirmă că mulțimea care constă din obiectele care se află fie în A, fie în A, este chiar mulțimea A, în timp ce 12) afirmă că mulțimea care constă

din acele obiecte care se află în A şi, de asemenea, fie în B, fie în C, este aceeaşi ca mulţimea formată din acele obiecte care se află fie în A şi B, fie în A şi C. Raţionamentul logic utilizat pentru demonstrarea acestor reguli asemănătoare poate fi ilustrat reprezentînd mulţimile A, B, C prin figuri aflate într-un plan, cu condiţia să avem grijă să considerăm toate posibilităţile referitoare la mulţimile implicate, în ceea ce priveşte elementele comune.

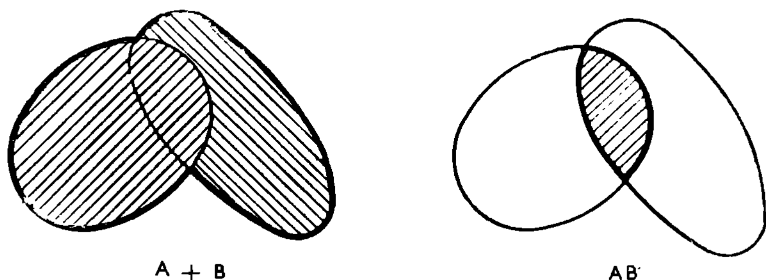


Fig. 26. Reuniunea şi intersecţia mulţimilor

Cititorul a observat că legile 6, 7, 8, 9 şi 12 sînt identice cu legile obişnuite ale comutativităţii, asociativităţii şi distributivităţii, cunoscute în algebră. Rezultă că toate regulile algebrei obişnuite a numerelor, care sînt consecinţe ale legilor comutativităţii, asociativităţii şi distributivităţii, sînt de asemenea valabile în algebra mulţimilor. Însă legile 10, 11, 13 nu au analoge numerice şi conferă algebrei mulţimilor o structură mai simplă decît cea a algebrei numerelor. De exemplu, teorema binomului din algebra obişnuită este înlocuită în algebra mulţimilor prin egalitatea

$$(A + B)^n = (A + B) \cdot (A + B) \dots (A + B) = A + B$$

care este o consecinţă a lui 11. Legile 14, 15 şi 17 indică faptul că proprietăţile lui $\emptyset$ şi I în raport cu reuniunea şi intersecţia mulţimilor sînt foarte asemănătoare cu proprietăţile numerelor 0 şi 1 faţă de adunarea şi înmulţirea obişnuită. Legea 16 nu are analog în algebra numerelor.

Rămîne să definim o altă operaţie în algebra mulţimilor. Fie A o submulţime a mulţimii universale I. Atunci, prin *complementara* lui A în I înţelegem mulţimea formată din toate obiectele lui I, care *nu* se află în A. Notăm această mulţime prin A'. Astfel, dacă I este mulţimea tuturor numerelor naturale, iar A este mulţimea numerelor prime, A' constă din 1 şi din numerele compuse. Operaţia A', care nu are un analog în algebra numerelor, posedă următoarele proprietăţi:

19) $A + A' = I$ 20) $AA' = \emptyset$
21) $\emptyset' = I$ 22) $I' = \emptyset$
23) $A'' = A$
24) Relația $A \subset B$ este echivalentă cu relația $B' \subset A'$.
25) $(A + B)' = A'B'$ 26) $(AB)' = A' + B'$.

Lăsăm din nou verificarea acestor legi pe seama cititorului.

Legile 1—26 formează baza algebrei mulțimilor. Ele posedă proprietatea remarcabilă de „dualitate", în sensul următor:

Dacă în una din legile 1—26 *înlocuim peste tot simbolurile*

$$\subset \text{ cu } \supset$$

$$\emptyset \text{ cu } I$$

$$+ \text{ cu } \cdot$$

(*în măsura în care ele apar*), *atunci rezultatul este din nou una din aceste legi.* De exemplu, legea 6 devine 7, 12 devine 13, 17 devine 16 etc. Rezultă că *oricărei teoreme care poate fi demonstrată pe baza legilor* 1—26 *îi corespunde alta, teorema „duală", obținută prin înlocuirile de mai sus.* Într-adevăr, dat fiind că demonstrația oricărei teoreme constă din aplicarea succesivă, la fiecare pas, a uneia din legile 1—26, aplicarea la fiecare pas a legii duale va da o demonstrație a teoremei duale. (Pentru o dualitate asemănătoare în geometrie, se poate consulta capitolul IV.)

2. Aplicație la logica matematică

Verificarea legilor algebrei mulțimilor se baza pe analiza semnificației logice a relației $A \subset B$ și a operațiilor $A + B$, AB și A'. Acum putem inversa acest proces și putem folosi legile 1—26 ca bază pentru o „algebră a logicii". Mai precis, acea parte a logicii care se referă la mulțimi sau, ceea ce este echivalent, la proprietățile sau *atributele* obiectelor, poate fi redusă la un sistem algebric formal, bazat pe legile 1—26. „Universul logic" definește mulțimea I; *orice proprietate sau atribut* $\mathfrak{A}$ *al obiectelor definește mulțimea* A *formată din toate obiectele din* I, *care au această proprietate.* Regulile de traducere a terminologiei logice uzuale în limbajul mulțimilor pot fi ilustrate prin următoarele exemple:

„*A* sau *B*"	$A + B$
„*A* și *B*"	AB
„Non *A*"	A'
„Nici *A* nici *B*"	$(A + B)'$, sau echivalent, $A'B'$
„Non (*A* și *B*)"	$(AB)'$, sau echivalent, $A' + B'$

„Toți A sînt B" sau „Dacă A, atunci B" sau „A implică B" — $A \subset B$
„Unii A sînt B" — $AB \neq \emptyset$
„Nici un A nu este B" — $AB = \emptyset$
„Unii A nu sînt B" — $AB' \neq \emptyset$
„Nu există nici un A" — $A = \emptyset$.

În termenii algebrei mulțimilor, silogismul „Barbara", care spune: „Dacă toți A sînt B, și toți B sînt C, atunci toți A sînt C", devine

3) Dacă $A \subset B$ și $B \subset C$, atunci $A \subset C$.

În mod similar, „legea contradicției", care spune: „Un obiect nu poate să posede o proprietate și să nu o posede", se scrie sub forma

(20) $$AA' = \emptyset,$$

în timp ce „legea terțiului exclus" care spune: „Un obiect trebuie fie să posede un atribut dat, fie să nu-l posede" devine

(19) $$A + A' = I.$$

Astfel, partea logicii, care este exprimabilă cu ajutorul simbolurilor $\subset$, $+$, $\cdot$, și $'$ poate fi tratată ca sistem algebric formal, supus legilor 1—26. Această fuziune a analizei logice a matematicii cu analiza matematică a logicii a creat o nouă disciplină, *logica matematică*, care se află în prezent într-un proces de dezvoltare puternică.

Din punctul de vedere al axiomaticii, este remarcabil faptul că propozițiile 1—26, împreună cu toate celelalte teoreme ale algebrei mulțimilor, pot fi deduse din următoarele trei ecuații:

(27) $$\begin{gathered} A + B = B + A \\ (A + B) + C = A + (B + C) \\ (A' + B')' + (A' + B)' = A. \end{gathered}$$

Rezultă că algebra mulțimilor poate fi construită ca teorie pur deductivă, ca și geometria euclidiană, pe baza acestor trei propoziții luate ca axiome. Făcînd aceasta, operația AB și relația de ordine $A \subset B$ se *definesc* cu ajutorul operațiilor $A + B$ și A'.

AB înseamnă mulțimea $(A' + B')'$, iar

$A \subset B$ înseamnă că $A + B = B$.

Un exemplu cu totul diferit de sistem matematic, care satisface toate legile formale ale algebrei mulțimilor, este dat de numerele 1, 2, 3, 5, 6, 10, 15, 30, în care $a + b$ înseamnă, prin definiție, cel mai mic multiplu comun al lui a și b, ab înseamnă cel mai mare divizor comun al lui a și b, $a \subset b$ în-

seamnă propoziția "a este un factor al lui b", și a' este numărul $30/a$. Existența acestor exemple a condus la studiul sistemelor algebrice generale, care satisfac legile 27). Aceste sisteme se numesc „algebre booleene" în memoria lui George Boole (1815—1864), matematician și logician englez, a cărui carte *An Investigation of the Laws of Thought* a apărut în 1854.

3. O aplicație la teoria probabilităților

Algebra mulțimilor clarifică foarte mult teoria probabilităților. Pentru a considera doar cel mai simplu caz, să ne imaginăm un experiment cu un număr finit de rezultate posibile, fiecare fiind presupus a fi „egal probabil". De exemplu, experimentul poate consta din extragerea unei cărți, la întîmplare, dintr-un pachet bine amestecat de 52 de cărți de joc. Dacă mulțimea rezultatelor posibile ale experimentului este notată cu I și dacă cu A notăm o submulțime a lui I, atunci probabilitatea ca rezultatul experimentului să aparțină submulțimii A este definită prin raportul

$$p(A) = \frac{\text{numărul elementelor din } A}{\text{numărul elementelor din } I}.$$

Dacă notăm numărul elementelor dintr-o mulțime A prin simbolul $n(A)$, atunci această definiție poate fi scrisă sub forma

$$p(A) = \frac{n(A)}{n(I)}. \tag{1}$$

În exemplul nostru, dacă A este submulțimea cupelor, atunci $n(A) = 13$, $n(I) = 52$, iar $p(A) = \frac{13}{52} = \frac{1}{4}$.

Noțiunile algebrei mulțimilor intervin în calculul probabilităților atunci cînd sînt cunoscute probabilitățile anumitor mulțimi și se cer probabilitățile altora. De exemplu, cunoscînd $p(A)$, $p(B)$ și $p(AB)$ putem calcula probabilitatea lui $p(A + B)$:

$$p(A + B) = p(A) + p(B) - p(AB). \tag{2}$$

Demonstrația este simplă. Avem

$$n(A + B) = n(A) + n(B) - n(AB),$$

deoarece elementele comune lui A și B, adică elementele din AB, vor fi numărate de două ori în suma $n(A) + n(B)$ și deci trebuie să scădem pe $n(AB)$ din această sumă, pentru a obține valoarea corectă a lui $n(A + B)$. Împărțind fiecare termen al acestei egalități prin $n(I)$, obținem egalitatea (2).

O formulă mai interesantă se obține considerînd trei submulțimi A, B, C ale lui I. Din (2) avem:

$$p(A + B + C) = p[(A + B) + C] = p(A + B) + p(C) - p[(A + B)C].$$

Din formula (12) a secţiunii precedente ştim că $(A + B)C = AC + BC$. Deci

$$p[(A + B)C] = p(AC + BC) = p(AC) + p(BC) - p(ABC).$$

Substituind această valoare a lui $p[(A + B)C]$ în egalitatea precedentă şi valoarea lui $p(A + B)$ dată de (2), obţinem formula dorită:

$$(3) \qquad p(A + B + C) = p(A) + p(B) + p(C) - p(AB) - p(AC) - p(BC) + p(ABC).$$

Ca exemplu, să considerăm experimentul următor. Cele trei cifre 1, 2, 3 sînt scrise la întîmplare. Care este probabilitatea ca cel puţin una din ele să ocupe locul pe care îl indică? Fie A mulţimea tuturor aranjărilor, în care cifra 1 se află pe primul loc, B mulţimea tuturor aranjărilor în care cifra 2 se află pe al doilea loc, iar C mulţimea tuturor aranjărilor în care cifra 3 se află pe al treilea loc. Atunci dorim să calculăm $p(A + B + C)$.

Este clar că

$$p(A) = p(B) = p(C) = \frac{2}{6} = \frac{1}{3};$$

deoarece dacă o cifră se află pe locul ei, rămîn două posibilităţi pentru celelalte cifre dintr-un total de $3 \cdot 2 \cdot 1 = 6$ aranjări posibile ale acestor trei cifre. Mai mult,

$$p(AB) = p(AC) = p(BC) = \frac{1}{6}$$

şi

$$p(ABC) = \frac{1}{6},$$

deoarece există un singur mod în care toate aceste trei cazuri pot avea loc. Din (3) rezultă că:

$$p(A + B + C) = 3 \cdot \frac{1}{3} - 3\left(\frac{1}{6}\right) + \frac{1}{6} = 1 - \frac{1}{2} + \frac{1}{6} = \frac{2}{3} = 0{,}6666\ldots$$

Exerciţiu: Găsiţi o formulă corespunzătoare pentru $p(A + B + C + D)$ şi aplicaţi-o pentru cazul a patru cifre. Probabilitatea corespunzătoare este $\frac{5}{8} = 0{,}6250$.

Formula generală pentru reuniunea a n submulţimi este

$$(4) \qquad p(A_1 + A_2 + \ldots + A_n) = \sum_1 p(A_i) - \sum_2 p(A_i A_j) + \sum_3 p(A_i A_j A_k) - \ldots \pm p(A_1 A_2 \ldots A_n),$$

în care simbolurile $\sum_1, \sum_2, \sum_3, \ldots, \sum_{n-1}$ indică adunarea combinaţiilor posibile ale mulţimilor $A_1, A_2, \ldots, A_n$, luate cîte una, două, trei, ..., $(n - 1)$. Această formulă poate fi stabilită prin inducţie matematică, tot aşa cum am dedus formula (3) din formula (2). Din (4)

rezultă cu ușurință că dacă scriem cele n cifre 1, 2, 3, ... n la întîmplare, probabilitatea ca cel puțin una din cifre să se afle pe locul pe care îl indică este egală cu

$$(5) \qquad p_n = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots \pm \frac{1}{n!},$$

unde ultimul termen este luat cu semnul plus sau minus, după cum n este impar sau par. În particular, pentru $n = 5$, probabilitatea este

$$p_5 = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!} = \frac{19}{30} = 0{,}63333\ \ldots.$$

Vom vedea în capitolul VIII că, atunci cînd n tinde către infinit, expresia

$$S_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots \pm \frac{1}{n!}$$

tinde spre limita 1/e, a cărei valoare cu cinci zecimale exacte este 0,36788. Deoarece din (5) rezultă că $p_n = 1 - S_n$, deducem că atunci cînd n tinde către infinit avem

$$p_n \to 1 - 1/e = 0{,}63212.$$

Capitolul III

CONSTRUCŢII GEOMETRICE. ALGEBRA CÎMPURILOR DE NUMERE

INTRODUCERE

Problemele de construcţie au fost întotdeauna un subiect de predilecţie în geometrie. Numai cu rigla şi cu compasul se pot face numeroase construcţii, după cum cititorul îşi aminteşte din şcoală: un segment sau un unghi pot fi împărţite în două părţi egale, se poate coborî perpendiculara dintr-un punct pe o dreaptă dată, se poate înscrie un hexagon regulat într-un cerc etc. În toate aceste probleme, rigla e folosită doar ca instrument de trasare a dreptei şi nu ca instrument de măsură sau de determinare a distanţelor. Restricţia tradiţională la riglă şi compas provine din antichitate, cu toate că grecii nu ezitau să folosească şi alte intrumente.

Una dintre problemele clasice celebre de construcţie este aşa-numita problemă de contact a lui Apollonius (aproximativ anul 200 î.e.n.), în care sînt date trei cercuri arbitrare într-un plan şi se cere un al patrulea cerc, tangent celor trei cercuri date. În particular, se admite ca unul sau mai multe dintre cercurile date să degenereze într-un punct sau o dreaptă (respectiv un „cerc" de rază zero sau „infinit"). De exemplu, se poate cere să se construiască un cerc tangent la două drepte date, care să treacă printr-un punct dat. În timp ce astfel de cazuri particulare pot fi tratate cu uşurinţă, problema generală este mult mai dificilă.

Dintre toate problemele de construcţie, aceea a construirii unui poligon regulat cu n laturi, folosind doar rigla şi compasul, prezintă poate cel mai mare interes. Pentru anumite valori ale lui n, de exemplu $n = 3, 4, 5, 6$, soluţia era cunoscută din antichitate şi formează o parte importantă a geometriei învăţate în şcoală. Însă pentru heptagonul regulat ($n = 7$) s-a dovedit că construcţia este imposibilă. Mai există trei probleme clasice, enunţate de greci, a căror soluţie s-a căutat în zadar: trisecţiunea unui unghi arbitrar dat, dublarea unui cub dat (adică găsirea muchiei unui cub al cărui volum să fie egal cu dublul volumului cubului cu muchie dată) şi cvadratura cercului (adică construirea unui pătrat care să aibă aceeaşi arie ca şi un cerc dat).

În toate aceste trei probleme, rigla și compasul sînt singurele instrumente permise.

Probleme nerezolvate de acest fel au dat naștere unora dintre dezvoltările noi și remarcabile ale matematicii, cînd, după secole de încercări zadarnice de rezolvare, s-a ivit bănuiala că aceste probleme sînt irevocabil nerezolvabile. Astfel, matematicienii au fost puși să cerceteze problema; *cum se poate demonstra că anumite probleme nu pot fi rezolvate?*

În algebră, problema rezolvării ecuațiilor de grad cinci și de grad superior a fost cea care a dus la acest nou mod de gîndire. În secolul al XVI-lea, matematicienii au învățat că ecuațiile algebrice de grad 3 sau 4 pot fi rezolvate printr-un proces asemănător cu metoda elementară de rezolvare a ecuațiilor pătratice. Toate aceste metode au următoarea caracteristică comună: soluțiile sau „rădăcinile" ecuației pot fi scrise ca expresii algebrice obținute din coeficienții ecuației printr-un șir de operații, fiecare fiind fie o operație rațională — adunare, scădere, înmulțire sau împărțire — fie extragerea unei rădicini — pătratice, cubice sau de ordinul patru. Se spune că ecuațiile algebrice de grad cel mult egal cu patru pot fi rezolvate „prin radicali" (de la cuvîntul *radix*, care înseamnă rădăcină în limba latină). Nimic nu părea mai natural decît extinderea acestui procedeu la ecuații de grad cinci sau grad superior, folosind firește, radicali, de ordin corespunzător. Toate aceste încercări au dat însă greș. Chiar matematicieni distinși din secolul al XVIII-lea s-au înșelat crezînd că au găsit soluția. Abia în primii ani ai secolului al XIX-lea, italianul Ruffini (1765—1822) și genialul matematician norvegian N.H. Abel (1802—1829) au conceput, ideea, revoluționară pe atunci, a demonstrării *imposibilității rezolvării prin radicali a ecuației algebrice generale de gradul n*. Trebuie să înțelegem limpede că nu se pune problema dacă orice ecuație algebrică de gradul *n posedă* soluții. Acest fapt a fost demonstrat de Gauss în teza sa de doctorat, în 1799, astfel încît nu există nici o îndoială în privința *existenței* rădăcinilor ecuației. Menționăm că aceste rădăcini pot fi găsite prin anumite procedee, cu orice precizie dorită. Arta rezolvării numerice a ecuațiilor, care are o foarte mare importanță practică, era admirabil elaborată. Dar problema lui Abel și Ruffini era cu totul deosebită: *se poate obține soluția doar cu ajutorul operațiilor raționale și al radicalilor?* Dorința de a obține o deplină claritate în această problemă a fost aceea care a inspirat dezvoltarea grandioasă a algebrei moderne și a teoriei grupurilor, ale căror baze au fost puse de Ruffini, Abel și Galois (1811—1832).

Problema demonstrării imposibilității unei anumite construcții geometrice dă unul dintre cele mai simple exemple ale acestei tendințe din algebră. Folosind noțiunile algebrice, vom fi în măsură să demonstrăm în acest capitol imposibilitatea trisecțiunii unghiului, a construirii heptagonului regulat sau a dublării cubului, doar cu ajutorul riglei și compasului. (Problema cvadraturii cercului este mult mai dificilă; cf. p.157). Punctul nostru de plecare nu va fi problema negativă a imposibilității anumitor construcții, ci mai curînd pro-

blema pozitivă: cum pot fi caracterizate complet toate problemele ale căror construcții pot fi făcute? După ce am răspuns la această întrebare, va fi ușor de arătat că problemele menționate mai sus nu fac parte din această categorie.

La vîrsta de 17 ani, Gauss a investigat problema construirii „p-goanelor" (poligoane cu p laturi) regulate, unde p este un număr prim. Pe atunci, construcția era cunoscută doar pentru $p = 3$ și $p = 5$. Gauss a descoperit că p-gonul regulat este construibil dacă și numai dacă numărul prim p este un „număr Fermat" de forma

$$p = 2^{2^n} + 1.$$

Primele numere Fermat sînt 3, 5, 17, 257, 65 537 (cf. p. 42). Tînărul Gauss a fost atît de copleșit de descoperirea sa, încît a renunțat imediat de a mai deveni filolog și s-a hotărît să-și consacre întreaga viață matematicii și aplicațiilor ei. El privea întotdeauna cu o deosebită mîndrie această primă mare descoperire a sa. După moarte, i-a fost ridicată o statuie de bronz la Göttingen, cu piedestalul în forma unui poligon regulat cu 17 laturi.

Cînd ne ocupăm de construcții geometrice, nu trebuie să uităm niciodată că nu se pune problema trasării figurilor în practică, cu un anumit grad de precizie, ci se pune problema dacă soluția poate fi găsită teoretic, folosind numai rigla și compasul, presupunînd că instrumentele sînt de o precizie perfectă. Ceea ce Gauss a demonstrat este faptul că construcțiile lui pot fi efectuate în principiu. Teoria lui nu se referă la efectuarea lor pe calea cea mai simplă, sau la găsirea artificiilor care pot fi folosite pentru a simplifica rezolvarea problemei. Aceasta este o problemă de importanță teoretică mult mai mică. Din punct de vedere practic, nici o construcție de acest fel nu ar da un rezultat satisfăcător în comparație cu acela obținut cu ajutorul unui raportor bun. Încapacitatea de a înțelege, așa cum trebuie, caracterul teoretic al problemei construcției geometrice, și încăpățînarea în refuzul de a lua cunoștință de fapte științifice bine stabilite, sînt cauza persistenței unei succesiuni nesfîrșite de trisectori de unghiuri și cvadratori de cercuri. Aceia care sînt capabili să înțeleagă matematica elementară ar putea profita din studiul acestui capitol.

Ar trebui să subliniem încă o dată că, din anumite privințe, noțiunea noastră de construcție geometrică pare artificială. Rigla și compasul sînt, desigur, cele mai simple instrumente de desen, dar restrîngerea la aceste instrumente nu este nicidecum inerentă geometriei. Matematicienii greci au recunoscut cu mult în urmă că anumite probleme — de exemplu aceea a dublării cubului — pot fi rezolvate dacă, de exemplu, este permisă folosirea unei rigle în formă de unghi drept; tot atît de ușor este să inventăm instrumente, altele decît compasul, cu ajutorul cărora putem trasa elipse, hiperbole, și curbe mai compli-

cate şi a căror folosire lărgeşte în mod considerabil domeniul figurilor construibile. În secţiunile următoare însă, vom adera la noţiunea clasică de construcţie geometrică, care foloseşte doar rigla şi compasul.

PARTEA I

DEMONSTRAŢIILE DE IMPOSIBILITATE ŞI ALGEBRA

§ 1. CONSTRUCŢII GEOMETRICE FUNDAMENTALE

1. Construirea cîmpurilor şi extragerea rădăcinii pătrate

Pentru a formula ideile noastre generale, vom începe prin examinarea cîtorva construcţii clasice. Cheia unei înţelegeri mai profunde se află în traducerea problemelor geometrice în limbajul algebric. Orice problemă de construcţie geometrică este de unul din următoarele tipuri: Se dă o anumită mulţime de segmente, a, b, c, ... şi se cer unul sau mai multe segmente x, y, ... Totdeauna este posibil să formulăm problemele în acest mod, chiar dacă la prima vedere ele au aspect cu totul deosebit. Segmentele cerute pot apare ca laturi ale unui triunghi care trebuie construit, ca raze ale unor cercuri, sau coordonate rectangulare ale anumitor puncte (cf. de exemplu, p. 154). Pentru a simplifica, vom presupune că se cere un singur segment x. Construcţia geometrică se reduce atunci la o problemă algebrică: mai întîi trebuie să găsim o relaţie între cantitatea cerută x şi cantităţile date a, b, c, ... ; apoi trebuie să găsim cantitatea necunoscută x prin rezolvarea acestei ecuaţii si, în sfîrşit, trebuie să determinăm dacă această soluţie poate fi obţinută prin procedee algebrice corespunzătoare construcţiilor cu rigla şi compasul. Principiul geometriei analitice, care constă în caracterizarea cantitativă a obiectelor geometrice cu ajutorul numerelor reale, bazată pe introducerea continuului numerelor reale, constituie fundamentul întregii teorii.

În primul rînd, observăm că unele dintre cele mai simple operaţii algebrice corespund unor construcţii geometrice elementare. Dacă sînt date două segmente cu lungimile a şi b (măsurate cu ajutorul unui segment „unitate" dat) atunci este foarte uşor de construit $a + b$, $a - b$, ra (unde r este un număr raţional oarecare), a/b, şi ab.

Pentru a construi pe $a + b$ (fig. 27), trasăm o dreaptă şi pe ea determinăm cu ajutorul compasului distanţele $OA = a$ şi $AB = b$. Atunci $OB = a + b$. În mod asemănător, pentru $a - b$ determinăm pe $OA = a$ şi pe $AB = b$, însă de data aceasta luăm pe AB în sens opus sensului lui OA. Atunci $OB = a - b$. Pentru a construi pe $3a$, adunăm pur şi simplu $a + a + a$; în mod asemănător construim pe pa, unde p este un întreg oarecare. Construim pe $a/3$ prin

artificiul următor (fig. 28): determinăm pe $OA = a$ pe o dreaptă și trasăm o a doua dreaptă prin O. Pe această dreaptă determinăm un segment arbitrar $OC = c$, și construim pe $OD = 3c$. Unim pe A cu D și ducem prin C o dreaptă paralelă cu AD, care intersectează pe OA în B. Triunghiurile OBC și OAD sînt asemenea; deci $OB/a = OB/OA = OC/OD = 1/3$ și $OB = a/3$. În același mod, putem construi pe a/q, unde q este un întreg oarecare. Efectuînd această operație asupra segmentului pa, putem construi în acest mod pe ra, unde $r = p/q$ este un număr rațional oarecare.

Pentru a construi pe a/b (fig. 29) determinăm pe $OB = b$ și $OA = a$ pe laturile unui unghi oarecare O, și pe OB luăm $OD = 1$. Prin D ducem o paralelă la AB, care intersectează pe OA în C. Atunci OC va avea lungimea a/b. Construcția lui ab este arătată în fig. 30, unde AD este paralela dusă prin A la BC.

Din aceste considerații rezultă că *operațiile algebrice „raționale"* — adunarea, scăderea, înmulțirea și împărțirea unor cantități cunoscute — *pot fi*

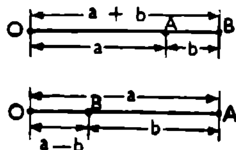


Fig. 27. Construirea lui $a + b$ *și* $a - b$

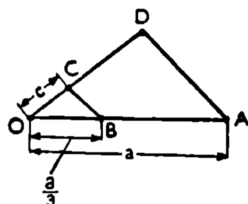


Fig. 28. Construirea lui a/3

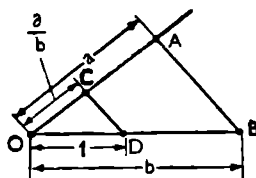


Fig. 29. Construirea lui a/b

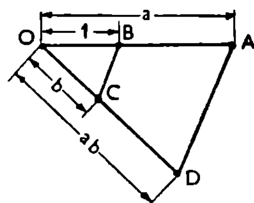


Fig. 30. Construirea lui ab

efectuate cu ajutorul unor construcții geometrice. Din orice segmente date, măsurate prin numerele reale $a, b, c, \ldots$, prin aplicarea succesivă a acestor construcții simple putem construi orice cantitate exprimabilă în funcție de $a,b,c, \ldots$ într-un mod rațional, adică prin aplicarea repetată a adunării, scăderii, înmulțirii, împărțirii. Totalitatea cantităților care pot fi obținute în

acest mod din a, b, c, ... constituie așa-numitul *cîmp de numere*, adică o mulțime de numere, astfel încît orice operații raționale aplicate la două sau mai multe elemente ale mulțimii dau tot un număr al mulțimii. Reamintim că numerele raționale, numerele reale și numerele complexe formează astfel de cîmpuri. În cazul de față, se spune că cîmpul este *generat* de numerele date a, b, c, ...

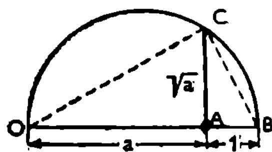


Fig. 31. Construirea lui $\sqrt{a}$

Construcția nouă, decisivă, care ne scoate din cîmpul obținut, este extragerea unei rădăcini pătrate: dacă este dat segmentul a, atunci $\sqrt{a}$ poate fi construit de asemenea, folosind doar rigla și compasul. Pe o dreaptă determinăm pe $OA = a$ și pe $AB = 1$ (fig. 31). Trasăm un cerc cu diametrul OB și construim perpendiculara la OB în punctul A, care intersectează cercul în C. Triunghiul OBC are un unghi drept în C, în virtutea teoremei din geometria elementară, care afirmă că un unghi înscris într-un semicerc este un unghi drept. Deci, $\sphericalangle OCA = \sphericalangle ABC$, triunghiurile dreptunghice OAC și CAB sînt asemenea și pentru $x = AC$ avem

$$\frac{a}{x} = \frac{x}{1}, \quad x^2 = a, \quad x = \sqrt{a}.$$

2. Poligoane regulate

Să considerăm acum cîteva probleme de construcție puțin mai complicate. Începem cu ***decagonul regulat***. Să presupunem că un decagon regulat este înscris într-un cerc de rază 1 (fig. 32) și să notăm cu x latura lui. Deoarece x va subîntinde un unghi de $36°$ în centrul cercului, celelalte două unghiuri ale triunghiului mare vor fi, ambele, de $72°$ și deci linia punctată care bisectează unghiul A, împarte triunghiul OAB în două triunghiuri isoscele, fiecare avînd laturile egale, egale cu x. Raza cercului este împărțită în acest mod în segmentele x și $1 - x$. Deoarece OAB este asemenea cu triunghiul isoscel mai mic, avem $1/x = x/(1 - x)$. Din această proporție obținem ecuația pătratică $x^2 + x - 1 = 0$, a cărei soluție este $x = (\sqrt{5} - 1)/2$. (Cealaltă soluție a ecuației este nesemnificativă, deoarece ea dă o valoare negativă pentru x.). De aici rezultă că x poate fi construit geometric. Avînd segmentul x, putem construi acum decagonul regulat, purtînd acest segment de zece ori ca

coardă în cerc. Pentagonul regulat poate fi construit acum unind vîrfurile decagonului regulat din două în două.

În loc de a construi pe $\sqrt{5}$ prin metoda din fig. 31, îl mai putem obține și ca ipotenuză a unui triunghi dreptunghic, ale cărui catete au lungimile 1 și 2. Atunci obținem pe x scăzînd unitatea de lungime din $\sqrt{5}$ și înjumătățind rezultatul.

Raportul $OB : AB$ din problema precedentă a fost numit raportul de aur, deoarece matematicienii greci considerau că un dreptunghi, ale cărui laturi se află în acest raport, este cel mai plăcut sub raport estetic. Valoarea lui este aproximativ egală cu 1,62.

Dintre toate poligoanele regulate, hexagonul este cel mai ușor de construit. Începem cu un cerc de rază r; lungimea laturii unui hexagon regulat înscris în acest cerc va fi atunci egală cu r. Hexagonul poate fi construit prin purtarea succesivă a unor coarde de lungime r, începînd dintr-un punct oarecare al cercului, pînă ce se obțin cele șase vîrfuri.

Din n-gonul regulat putem obține $2n$-gonul regulat împărțind în două părți egale arcul subîntins pe cercul circumscris, de fiecare latură a n-gonului, și folosind punctele suplimentare astfel obținute, ca și vîrfurile inițiale, ca vîrfuri ale $2n$-gonului cerut. Începînd cu diametrul unui cerc (un „2-gon"), putem construi de aceea 4, 8, 16, ..., 2^n-gonul. În mod similar, putem obține

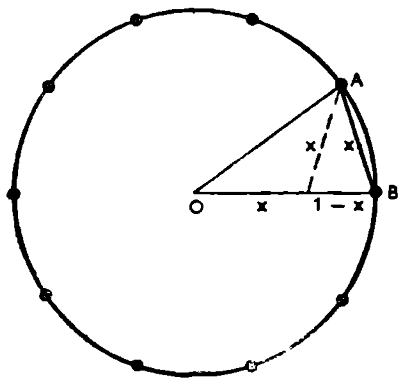


Fig. 32. Decagonul regulat

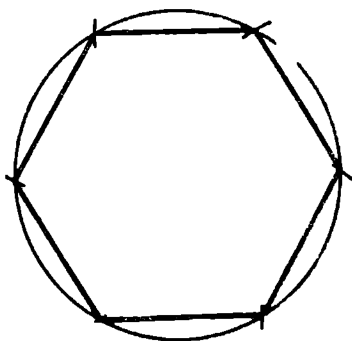


Fig. 33. Hexagonul regulat

12, 24, 48-gonul etc. pornind de la hexagon și 20, 40-gonul etc. pornind de la decagon.

Dacă s_n este lungimea laturii n-gonului regulat înscris în cercul unitate (cercul de rază 1), atunci latura $2n$-gonului are lungimea

$$s_{2n} = \sqrt{2 - \sqrt{4 - s_n^2}}.$$

Acest lucru poate fi demonstrat în felul următor: în fig. 34, s_n este egal cu $DE=2DC$, s_{2n} este egal cu DB, iar AB este egal cu 2. Aria triunghiului dreptunghic ABD este dată de $\frac{1}{2} BD \cdot AD$ și de $\frac{1}{2} AB \cdot CD$. Deoarece $AD = \sqrt{AB^2 - DB^2}$, făcînd substituțiile $AB = 2$, $BD = s_{2n}$, $CD = \frac{1}{2} s_n$ și egalînd cele două expresii ale ariei, găsim

$$s_n = s_{2n} \sqrt{4 - s_{2n}^2} \text{ sau } s_n^2 = s_{2n}^2(4 - s_{2n}^2).$$

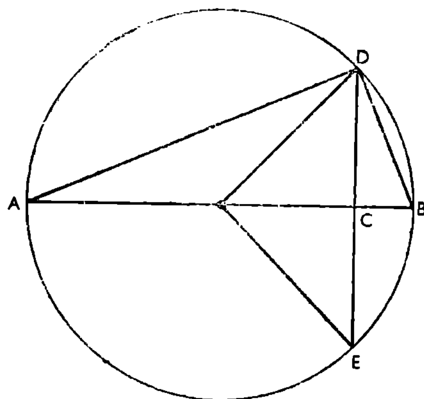


Fig. 34.

Rezolvînd această ecuație pătratică în raport cu $x = s_{2n}^2$ și observînd că x trebuie să fie mai mic decît 2, găsim cu ușurință formula dată mai sus.

Din această formulă și din faptul că s_4 (latura pătratului) este egală cu $\sqrt{2}$, rezultă că

$$s_8 = \sqrt{2 - \sqrt{2}}, \quad s_{16} = \sqrt{2 - \sqrt{2 + \sqrt{2}}},$$

$$s_{32} = \sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2}}}}, \text{ etc.}$$

Ca formulă generală obținem pentru $n > 2$,

$$s_{2^n} = \sqrt{2 - \sqrt{2 + \sqrt{2 + \ldots + \sqrt{2}}}}$$

cu $n - 1$ radicali suprapuși. Perimetrul 2^n-gonului înscris în cerc este egal cu $2^n s_{2^n}$. Cînd n tinde către infinit, 2^n-gonul tinde către cerc. Deci $2^n s_{2^n}$ tinde spre lungimea circumferinței cercului

unitate, care prin definiție este egală cu 2π. Astfel, înlocuind pe $n - 1$ cu m și simplificînd cu factorul 2, obținem formula:

$$\underbrace{2^m \sqrt{2 - \sqrt{2 + \sqrt{2 + \ldots + \sqrt{2}}}}}_{m \text{ rădăcini pătrate.}} \rightarrow \pi \text{ cînd } m \rightarrow \infty$$

Exercițiu: Deoarece $2^m \rightarrow \infty$, demonstrați ca o consecință că

$$\underbrace{\sqrt{2 + \sqrt{2 + \ldots + \sqrt{2}}}}_{n \text{ rădăcini pătrate}} \rightarrow 2 \text{ cînd } n \rightarrow \infty$$

Rezultatele obținute pînă acum indică următorul fapt caracteristic: ***Laturile 2^n-gonului, ale $5 \cdot 2^n$-gonului și ale $3 \cdot 2^n$-gonului pot fi obținute, prin adunări, scăderi, înmulțiri, împărțiri și extrageri de rădăcini pătrate.***

*3. Problema lui Apollonius

O altă problemă de construcție, care devine foarte simplă din punct de vedere algebric, este celebra problemă de contact a lui Apollonius, pe care am mai menționat-o. În contextul nostru, nu este necesar să găsim o construcție deosebit de elegantă. Ceea ce contează aici este faptul, că, în principiu, problema poate fi rezolvată doar cu ajutorul riglei și al compasului. Vom da o scurtă indicație asupra demonstrației, amînînd problema unei metode mai elegante de construcție, pînă la p. 178.

Fie (x_1, y_1), (x_2, y_2) și (x_3, y_3) coordonatele centrelor cercurilor date, respectiv de raze r_1, r_2 și r_3. Să notăm cu (x, y) centrul cercului căutat și cu r raza lui. Atunci condiția ca cercul căutat să fie tangent celor trei cercuri date se obține observînd că distanța dintre centrele a două cercuri tangente este egală cu suma sau diferența razelor, după cum cercurile sînt tangente exterior sau interior. Aceasta ne duce la ecuațiile

(1) $$(x - x_1)^2 + (y - y_1)^2 - (r \pm r_1)^2 = 0,$$

(2) $$(x - x_2)^2 + (y - y_2)^2 - (r \pm r_2)^2 = 0,$$

(3) $$(x - x_3)^2 + (y - y_3)^2 - (r \pm r_3)^2 = 0,$$

sau

(1a) $$x^2 + y^2 - r^2 - 2xx_1 - 2yy_1 \pm 2rr_1 + x_1^2 + y_1^2 - r_1 = 0$$

etc. Semnul plus sau minus se alege în fiecare din aceste ecuații, după cum cercurile sînt tangente interior sau exterior (fig. 35). Ecuațiile (1), (2) și (3)

sînt trei ecuații pătratice cu trei necunoscute x, y, r, cu proprietatea că termenii de gradul doi sînt aceiași în fiecare ecuație, după cum se vede din forma dezvoltată (1a). Deci, scăzînd ecuația (2) din (1), obținem o ecuație liniară în x, y, r:

$$ax + by + cr = d, \tag{4}$$

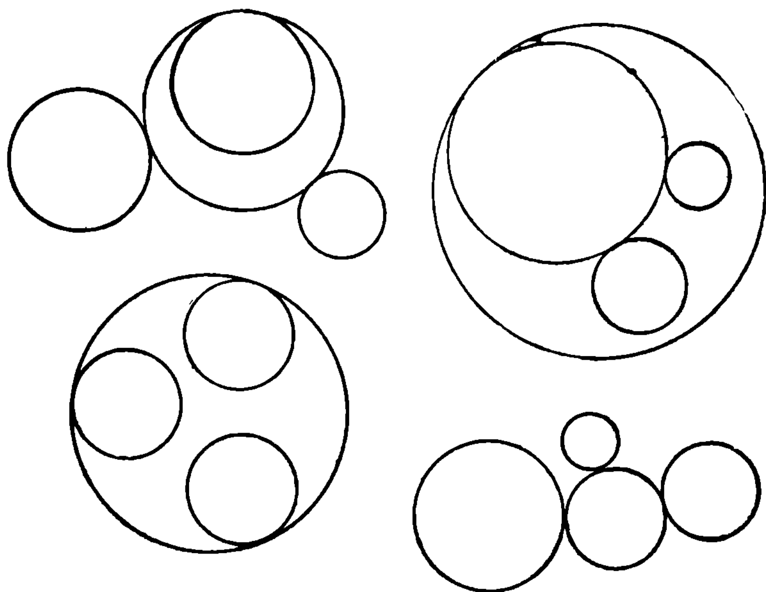


Fig. 35. Cercurile lui Apollonius

unde $a = 2(x_2 - x_1)$ etc. În mod analog, scăzînd ecuația (3) din (1) obținem o altă ecuație liniară

$$a'x + b'y + c'r = d'. \tag{5}$$

Rezolvînd ecuațiile (4) și (5) în raport cu x și y și substituind valorile găsite în ecuația (1), obținem o ecuație pătratică în raport cu r, care poate fi rezolvată prin operații raționale și prin extragerea unei rădăcini pătrate (cf. p.107). În general, vom avea două soluții pentru această ecuație, dintre care numai una va fi pozitivă. După ce găsim pe r din această ecuație, obținem pe x și pe y din cele două ecuații liniare (4) și (5). Cercul cu centrul (x, y) și de rază r va fi tangent celor trei cercuri date. În procesul rezolvării am folosit doar operații raționale și extrageri de rădăcini pătrate. Rezultă că r, x și y pot fi construite doar cu ajutorul riglei și compasului.

În general vom avea opt soluții ale problemei lui Apollonius, corespunzătoare celor $2 \cdot 2 \cdot 2 = 8$ combinări posibile ale semnelor $+$ și $-$ în ecuațiile (1), (2), și (3). Aceste alegeri corespund condițiilor ca cercurile cerute să fie tangente exterior sau interior fiecăruia din cele trei cercuri date. Se poate întîmpla ca procedeul nostru algebric să nu ne dea valori reale pentru x, y și r. Acest lucru se va întîmpla, de exemplu, dacă cele trei cercuri date sînt concentrice, astfel încît nu există nici o soluție a problemei geometrice. De asemenea, ne putem aștepta la ,,degenerări'' ale soluției, ca în cazul în care cele trei cercuri date degenerează în trei puncte coliniare. Atunci, cercul lui Apollonius degenerează în aceeași dreaptă. Nu vom discuta amănunțit aceste posibilități; un cititor cu oarecare experiență algebrică va fi capabil să completeze această analiză.

§ 2. NUMERE CONSTRUIBILE ȘI CÎMPURI DE NUMERE

1. Teoria generală

Discuția noastră precedentă indică fundamentul algebric general al construcțiilor geometrice. Orice construcție cu rigla și compasul constă dintr-un șir de pași, fiecare fiind de unul din următoarele tipuri: 1) trasarea unei drepte prin două puncte, 2) aflarea intersecției a două drepte, 3) trasarea unui cerc de rază dată, cu centrul într-un punct dat, 4) aflarea punctelor de intersecție ale unui cerc cu un alt cerc sau cu o dreaptă. Un element (punct, dreaptă sau cerc) se consideră a fi cunoscut, dacă a fost dat de la început, sau dacă a fost construit într-unul din pașii precedenți. Pentru o analiză teoretică putem raporta întreaga construcție la un sistem de coordonate x, y (cf. p. 89). Elementele date vor fi reprezentate atunci prin puncte sau segmente în planul x, y. Dacă de la început este dat un singur segment, atunci îl putem lua ca unitate de lungime, ceea ce fixează punctul $x = 1, y = 0$. Uneori apar elemente ,,arbitrare'': se trasează drepte arbitrare, se aleg puncte sau raze arbitrare. (Un exemplu de element arbitrar de acest fel apare la construcția mijlocului unui segment; trasăm două cercuri de raze egale, dar arbitrare, cu centrul în fiecare extremitate a segmentului și unim punctele lor de intersecție.) În astfel de cazuri, putem alege elementul așa încît el să fie rațional; adică putem alege puncte arbitrare cu coordonate raționale x, y, drepte arbitrare $ax + by + c = 0$ cu coeficienți raționali a, b, c, cercuri arbitrare, ale căror centre au coordonate raționale și care au raze raționale. Vom face o astfel de alegere a elementelor arbitrare raționale peste tot; dacă elementele sînt într-adevăr arbitrare, această restricție nu poate influența rezultatul unei construcții.

Pentru simplificare, vom presupune în discuția care urmează că, de la început, este dat un singur element, unitatea de lungime 1. Atunci, în concordanță cu § 1, putem construi cu ajutorul riglei și compasului toate numerele

care pot fi obținute din unitate prin procesele raționale de adunare, scădere, înmulțire și împărțire, adică toate numerele raționale r/s, unde r și s sînt întregi. Sistemul numerelor raționale este „închis" în raport cu operațiile raționale; adică suma, diferența, produsul sau cîtul a două numere raționale — excluzînd împărțirea cu zero, ca de obicei — este din nou un număr rațional. Orice mulțime de numere, care are această proprietate de închidere în raport cu cele patru operații raționale, se numește *cîmp de numere.*

Exercițiu: Arătați că orice cîmp conține toate numerele raționale. (Indicație: Dacă $a \neq 0$ este un număr al cîmpului F, atunci $a/a = 1$ aparține lui F și din 1 putem obține orice număr rațional prin operații raționale.)

Pornind de la unitate, putem construi astfel întregul cîmp al numerelor raționale și deci toate punctele raționale (adică punctele care au ambele coordonate raționale) din planul x, y. Putem obține noi numere, iraționale, folosind compasul pentru a construi, de pildă, numărul $\sqrt{2}$, care, după cîte știm din capitolul II, § 2, nu se află în cîmpul rațional. După ce am construit pe $\sqrt{2}$, putem găsi prin construcțiile „raționale" din § 1, toate numerele de forma

$$(1) \qquad a + b\sqrt{2},$$

unde a, b sînt numere raționale, și de aceea și ele sînt construibile. De asemenea, putem construi toate numerele de forma

$$\frac{a + b\sqrt{2}}{c + d\sqrt{2}} \text{ sau } (a + b\sqrt{2})(c + d\sqrt{2}),$$

unde a, b, c, d sînt raționale. Aceste numere pot fi scrise însă întotdeauna sub forma (1), deoarece avem

$$\frac{a + b\sqrt{2}}{c + d\sqrt{2}} = \frac{a + b\sqrt{2}}{c + d\sqrt{2}} \cdot \frac{c - d\sqrt{2}}{c - d\sqrt{2}} = \frac{ac - 2bd}{c^2 - 2d^2} + \frac{bc - ad}{c^2 - 2d^2}\sqrt{2} = p + q\sqrt{2},$$

unde p, q sînt raționale. (Numitorul $c^2 - 2d^2$ nu poate fi nul, deoarece dacă $c^2 - 2d^2 = 0$, atunci $\sqrt{2} = c/d$, contrar faptului că $\sqrt{2}$ este irațional.) De asemenea,

$$(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (bc + ad)\sqrt{2} = r + s\sqrt{2},$$

unde r, s sînt raționale. Prin urmare, tot ceea ce obținem prin construirea lui $\sqrt{2}$ este mulțimea numerelor de forma (1), cu numerele raționale, a, b, arbitrare.

Exerciții: Din $p = 1 + \sqrt{2}$, $q = 2 - \sqrt{2}$, $r = -3 + \sqrt{2}$ obțineți numerele

$$\frac{p}{q},\ p + p^2,\ \left(p - p^2\right)\frac{q}{r},\ \frac{pqr}{1 + r^2},\ \frac{p + q^r}{q + pr^2},$$

sub forma (1).

Aceste numere (1) formează din nou un cîmp, după cum rezultă din discuția precedentă. (Faptul că suma și diferența a două numere de forma (1) este tot un număr de forma (1) este evident.) Acest cîmp este mai cuprinzător decît cîmpul rațional, care este o parte, sau un *subcîmp* al său. Însă, desigur, este mai mic decît cîmpul *tuturor* numerelor reale. Să notăm cîmpul rațional cu F_0 și noul cîmp al numerelor de forma (1) cu F_1. Faptul că orice număr din „cîmpul extins" F_1 poate fi construit a fost stabilit. Putem extinde acum domeniul construcțiilor noastre luînd, de exemplu, un număr din F_1, ca de pildă $k = 1 + \sqrt{2}$, extrăgînd rădăcina pătrată și obținînd în acest mod numărul construibil

$$\sqrt{1 + \sqrt{2}} = \sqrt{k};$$

o dată cu el, conform § 1, obținem cîmpul format din toate numerele de forma

(2) $$p + q\sqrt{k},$$

unde acum p și q pot fi numere arbitrare din F_1, adică numere de forma $a + b\sqrt{2}$, unde a, b se află în F_0, adică sînt raționale.

Exerciții: Reprezentați numerele

$$(\sqrt{k})^3,\ \frac{1 + (\sqrt{k})^2}{1 + \sqrt{k}},\ \frac{\sqrt{2}\sqrt{k} + \frac{1}{\sqrt{2}}}{(\sqrt{k})^3 - 3},\ \frac{(1 + \sqrt{k})(2 - \sqrt{k})\left(\sqrt{2} + \frac{1}{\sqrt{k}}\right)}{1 + \sqrt{2}k},$$

sub forma (2).

Toate aceste numere au fost construite pe baza ipotezei că, de la început, a fost dat un singur segment. Dacă sînt date două segmente, putem alege pe unul din ele ca unitate de lungime. Să presupunem că, față de această unitate, lungimea celuilalt segment este egală cu α. Atunci putem construi cîmpul G format din toate numerele de forma

$$\frac{a_m\alpha^m + a_{m-1}\alpha^{m-1} + \ldots + a_1\alpha + a_0}{b_n\alpha^n + b_{n-1}\alpha^{n-1} + \ldots + b_1\alpha + b_0},$$

unde numerele $a_0, \ldots, a_m$ și $b_0, \ldots, b_n$ sînt raționale, iar m și n sînt întregi pozitivi arbitrari.

Exercițiu: Dacă sînt date două segmente de lungimi 1 și α, indicați construcțiile efective necesare pentru obținerea numerelor $1 + \alpha + \alpha^2$, $(1 + \alpha)/(1 - \alpha)$, α^3.

Să presupunem acum, mai general, că sîntem în stare să construim toate numerele unui cîmp numeric F. Vom arăta că *folosind doar rigla nu vom ieşi niciodată din cîmpul* F. Ecuaţia dreptei care trece prin punctele de coordonate a_1, b_1 şi a_2, b_2 luate din cîmpul F, este $(b_1 - b_2)x + (a_2 - a_1)y + (a_1b_2 - a_2b_1) = 0$ (cf. p. 511); coeficienţii ei sînt expresii raţionale formate cu numere din F, şi deci, din definiţia cîmpului, sînt şi ei în F. Mai mult, dacă avem două drepte $\alpha x + \beta y - \gamma = 0$ şi $\alpha' x + \beta' y - \gamma' = 0$, cu coeficienţi în F, atunci coordonatele punctului lor de intersecţie, aflate prin rezolvarea sistemului corespunzător de ecuaţii, sînt $x = \dfrac{\gamma\beta' - \beta\gamma'}{\alpha\beta' - \beta\alpha'}$, $y = \dfrac{\alpha\gamma' - \gamma\alpha'}{\alpha\beta' - \beta\alpha'}$. Deoarece şi acestea sînt numere din F, este limpede că folosind doar rigla nu putem ieşi din cîmpul F.

Exerciţii. Dreptele $x + \sqrt{2}\,y - 1 = 0$, $2x - y + \sqrt{2} = 0$, au coeficienţii în cîmpul (1). Calculaţi coordonatele punctului lor de intersecţie şi verificaţi faptul că ele sînt de forma (1). Uniţi punctele $(1, \sqrt{2})$ şi $(\sqrt{2}, 1 - \sqrt{2})$ printr-o dreaptă $ax + by + c = 0$ şi verificaţi faptul că coeficienţii sînt de forma (1). Faceţi acelaşi lucru în raport cu cîmpul (2) pentru dreptele $\sqrt{1 + \sqrt{2}}\,x + \sqrt{2}\,y = 1$, $(1 + \sqrt{2})\,x - y = 1 - \sqrt{1 + \sqrt{2}}$, şi punctele $(\sqrt{2}, -1)$, $(1 + \sqrt{2}, \sqrt{1 + \sqrt{2}})$.

Putem ieşi din limitele lui F numai prin folosirea compasului. În acest scop, alegem un element k al lui F, astfel încît $\sqrt{k}$ să nu se afle în F. Atunci putem construi pe $\sqrt{k}$, şi, de aceea, putem construi toate numerele de forma

$$a + b\sqrt{k}, \tag{3}$$

unde a şi b sînt raţionale, sau chiar elemente arbitrare ale lui F. Suma şi diferenţa a două numere de forma $a + b\sqrt{k}$ şi $c + d\sqrt{k}$, produsul lor $(a + b\sqrt{k})(a + d\sqrt{k}) = (ac + kbd) + (ad + bc)\sqrt{k}$ şi cîtul lor

$$\frac{a + b\sqrt{k}}{c + d\sqrt{k}} = \frac{(a + b\sqrt{k})(c - d\sqrt{k})}{c^2 - kd^2} = \frac{ac - kbd}{c^2 - kd^2} + \frac{bc - ad}{c^2 - kd^2}\sqrt{k},$$

sînt tot de forma $p + q\sqrt{k}$, unde p şi q sînt numere din F. (Numitorul $c^2 - kd^2$ nu se poate anula, decît dacă c şi d sînt ambii nuli; într-adevăr, în caz contrar, am avea $\sqrt{k} = c/d$, număr din F, contrar ipotezei după care $\sqrt{k}$ nu se află în F.) Deci, mulţimea numerelor de forma $a + b\sqrt{k}$ formează un cîmp F'. Cîmpul F' conţine cîmpul iniţial F, deoarece putem alege în particular $b = 0$. F' se numeşte o *extindere* a cîmpului F, iar F se numeşte *subcîmp* al cîmpului F'.

Ca exemplu, fie F cîmpul format din numerele $a + b\sqrt{2}$, unde a şi b sînt numere raţionale şi să luăm $k = \sqrt{2}$. Atunci numerele extinderii F' sînt reprezentate de $p+q\sqrt[4]{2}$, unde p şi q se află în F, $p = a + b\sqrt{2}$, $q=a'+b'\sqrt{2}$, cu a, b, a', b' raţionale. Orice număr din F' poate fi adus la această formă. De exemplu:

$$\frac{1}{\sqrt{2}+\sqrt[4]{2}} = \frac{\sqrt{2}-\sqrt[4]{2}}{(\sqrt{2}+\sqrt[4]{2})(\sqrt{2}-\sqrt[4]{2})} = \frac{\sqrt{2}-\sqrt[4]{2}}{2-\sqrt{2}} = \frac{\sqrt{2}}{2-\sqrt{2}} - \frac{\sqrt[4]{2}}{2-\sqrt{2}} =$$

$$= \frac{\sqrt{2}(2+\sqrt{2})}{4-2} - \frac{(2+\sqrt{2})}{4-2}\sqrt[4]{2} = (1+\sqrt{2}) - \left(1+\frac{1}{2}\sqrt{2}\right)\sqrt[4]{2}.$$

Exerciţiu: Fie F cîmpul $p + q\sqrt{2+\sqrt{2}}$, unde p şi q sînt de forma $a + b\sqrt{2}$, cu a şi b raţionale. Reprezentaţi numărul

$$\frac{1+\sqrt{2+\sqrt{2}}}{2-3\sqrt{2+\sqrt{2}}}$$

sub această formă.

Am văzut că dacă pornim de la orice cîmp F de numere construibile care conţine numărul k, atunci prin folosirea riglei şi printr-o singură aplicare a compasului, putem construi pe $\sqrt{k}$ şi deci orice număr de forma $a + b\sqrt{k}$, unde a şi b sînt numere din F.

Să arătăm acum că reciproc, printr-o singură aplicare a compasului, putem obţine *numai* numere de această formă. Într-adevăr ceea ce realizează compasul într-o construcţie este tocmai definirea punctelor (sau a coordonatelor lor) ca puncte ale intersecţiei unui cerc cu o dreaptă sau cu un cerc. Un cerc de centru ξ, η şi de rază r are ecuaţia $(x - \xi)^2 + (y - \eta)^2 = r^2$; deci, dacă ξ, η, r sînt în F, ecuaţia cercului poate fi scrisă sub forma

$$x^2 + y^2 + 2\alpha x + 2\beta y + \gamma = 0,$$

cu coeficienţii α, β, γ, din F. O dreaptă

$$ax + by + c = 0,$$

care uneşte două puncte, ale căror coordonate se află în F, are coeficienţii a, b, c în F, după cum am văzut la p. 147. Eliminînd pe y din acest sistem, obţinem pentru abscisa unui punct de intersecţie a cercului cu dreapta o ecuaţie pătratică de forma

$$Ax^2 + Bx + C = 0,$$

cu coeficienți A, B, C în F (în mod explicit $A = a^2 + b^2$, $B = 2(ac + b^2\alpha - ab\beta)$, $C = c^2 - 2bc\beta + b^2\gamma$). Soluția este dată de formula

$$x = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A},$$

care este de forma $p + q\sqrt{k}$, cu p, q, k, din F. O formulă asemănătoare este adevărată pentru ordonata unui punct de intersecție. Dacă însă avem două cercuri

$$x^2 + y^2 + 2\alpha x + 2\beta y + \gamma = 0,$$

$$x^2 + y^2 + 2\alpha' x + 2\beta' y + \gamma' = 0,$$

atunci, scăzînd a doua ecuație din prima, obținem ecuația liniară

$$2(\alpha - \alpha')x + 2(\beta - \beta')y + (\gamma - \gamma') = 0,$$

care poate fi rezolvată împreună cu ecuația primului cerc, ca și mai înainte. În ambele cazuri, construcția furnizează cele două coordonate ale noilor puncte și aceste noi cantități sînt de forma $p + q\sqrt{k}$, cu p, q, k din F. În particular, desigur, $\sqrt{k}$ poate aparține el însuși lui F, ca de pildă cînd $k = 4$. Atunci, construcția nu dă esențial nimic nou și rămînem în F. Dar, în general, lucrurile nu se vor petrece în acest fel.

Exerciții: Considerați cercul de rază $2\sqrt{2}$, cu centrul în origine și dreapta care trece prin punctele $(1/2, 0)$, $(4\sqrt{2}, \sqrt{2})$. Găsiți cîmpul F' determinat de coordonatele punctelor de intersecție ale cercului și ale dreptei. Faceți același lucru pentru intersecția cercului dat cu cercul de rază $\sqrt{2}/2$ și de centru $(0, 2\sqrt{2})$.

Rezumînd, dacă de la început sînt date anumite cantități, atunci doar cu ajutorul riglei putem construi toate cantitățile din cîmpul F, generat prin operații raționale efectuate cu cantitățile date. Folosind compasul, putem extinde apoi cîmpul F al cantităților construibile, obținînd o extindere a cîmpului dat, mai largă, alegînd un număr k din F, extrăgînd rădăcina pătrată din k, și construind cîmpul F' format din numerele $a + b\sqrt{k}$, unde a și b se află în F; F se numește subcîmp al cîmpului F'; toate cantitățile din F se află, de asemenea, în F', deoarece în expresia $a + b\sqrt{k}$ putem alege $b = 0$. (Presupunem că $\sqrt{k}$ este un nou număr, care nu se află în F, deoarece în caz contrar procesul adjuncționării lui $\sqrt{k}$ nu ne-ar da nimic nou și F' ar fi identic cu F.) Am arătat că orice pas efectuat într-o construcție geometrică (trasarea unei drepte prin două puncte cunoscute, trasarea unui cerc cu centrul cunoscut și de raza dată, sau intersectarea a două drepte sau cercuri cunoscute) va produce fie noi cantități, aflate în cîmpul despre care știm că este format

din numere construibile, sau prin construirea unei rădăcini pătrate, va produce o nouă extindere a cîmpului numerelor construibile.

Mulţimea tuturor numerelor construibile poate fi descrisă acum în mod precis. Începem cu un cîmp dat F_0, definit prin cantităţile date de la început, de pildă cîmpul numerelor raţionale, dacă este dat un singur segment, luat ca unitate de lungime. Apoi, prin ,,adjuncţionarea'' lui $\sqrt{k_0}$, unde k_0 se află în F_0, dar $\sqrt{k_0}$ nu se află în F_0, construim o extindere F_1 a numerelor construibile, formată din toate numerele de forma $a_0 + b_0\sqrt{k_0}$, unde a_0 şi b_0 pot fi numere oarecare din F_0. Atunci F_2, o nouă extindere a cîmpului F_1, este definită de numerele $a_1 + b_1\sqrt{k_1}$, unde a_1 şi b_1 sînt numere arbitrare din F_1, iar k_1 este un număr din F_1, a cărui rădăcină pătrată nu se află în F_1. Repetînd acest *procedeu*, vom obţine un cîmp F_n, după ,,adjuncţionarea'' a n rădăcini pătrate. *Numerele construibile sînt acelea şi numai acelea, care pot fi obţinute printr-un astfel de şir de extinderi; adică, acelea care se află într-un cîmp* F_n *de tipul descris.* Mărimea numărului n de extinderi necesare nu are importanţă; într-o oarecare măsură, el măsoară gradul de complexitate a problemei.

Exemplul următor ilustrează procedeul. Dorim să obţinem numărul

$$\sqrt{6} + \sqrt{\sqrt{\sqrt{1+\sqrt{2}} + \sqrt{3}} + 5}.$$

Fie F_0 cîmpul raţional. Punînd $k_0 = 2$, obţinem cîmpul F_1, care conţine numărul $1 + \sqrt{2}$. Acum luăm $k_1 = 1 + \sqrt{2}$ şi $k_2 = 3$. De fapt, 3 se află în cîmpul iniţial F_0 şi deci *a fortiori* în cîmpul F_2, astfel încît este permis să luăm $k_2 = 3$. Apoi luăm $k_3 = \sqrt{1+\sqrt{2}} + \sqrt{3}$ şi, în sfîrşit, $k_4 = \sqrt{\sqrt{1+\sqrt{2}}+\sqrt{3}} + 5$. Cîmpul F_5 astfel construit conţine numărul dorit, pentru că $\sqrt{6}$ se află în F_5, dat fiind că $\sqrt{2}$ şi $\sqrt{3}$, şi deci produsul lor, se află în F_3, deci se află în F_5.

Exerciţii: Verificaţi că, pornind de la cîmpul raţional, latura 2^m-gonului regulat (cf. p. 141) este un număr construibil, pentru $n = m - 1$. Determinaţi şirul extinderilor succesive. Faceţi acelaşi lucru pentru numerele

$$\sqrt{1+\sqrt{2}+\sqrt{3}+\sqrt{5}},\quad (\sqrt{5}+\sqrt{11})/(1+\sqrt{7-\sqrt{3}}),$$
$$(\sqrt{2+\sqrt{3}})(\sqrt[8]{2}+\sqrt{1+\sqrt{2+\sqrt{5}}+\sqrt{3-\sqrt{7}}}).$$

2. Toate numerele construibile sînt algebrice

Dacă cîmpul iniţial F_0 este cîmpul raţional generat de un singur segment, atunci toate numerele construibile vor fi algebrice. (Pentru definirea numerelor algebrice cf. p. 120). Numerele cîmpului F_1 sînt rădăcinile unor ecuaţii pătratice, acelea ale lui F_2 sînt rădăcinile unor ecuaţii

de gradul patru şi, în general, numerele lui F_k sînt rădăcinile unor ecuaţii de grad 2^k cu coeficienţi raţionali. Pentru a demonstra acest lucru, pentru un cîmp F_2, putem considera mai întîi, ca exemplu, pe $x = \sqrt{2} + \sqrt{3 + \sqrt{2}}$. Avem $(x - \sqrt{2})^2 = 3 + \sqrt{2}$, $x^2 + 2 - 2\sqrt{2}x = 3 + \sqrt{2}$, sau $x^2 - 1 = \sqrt{2}(2x + 1)$, o ecuaţie pătratică cu coeficienţii într-un cîmp F_1. Ridicînd la pătrat, obţinem, în sfîrşit,

$$(x^2 - 1)^2 = 2(2x + 1)^2,$$

care este o ecuaţie de gradul patru, cu coeficienţi raţionali.

În general, orice număr dintr-un cîmp F_2 este de forma:

$$x = p + q\sqrt{w}, \tag{4}$$

unde p, q, w se află într-un cîmp F_1, şi deci sînt de forma $p = a + b\sqrt{s}$, $q = c + d\sqrt{s}$, $w = e + f\sqrt{s}$, unde a, b, c, d, e, f, s sînt numere raţionale. Din (4) rezultă

$$x^2 - 2px + p^2 = q^2w$$

unde toţi coeficienţii se află într-un cîmp F_1, generat de $\sqrt{s}$. Deci această ecuaţie poate fi scrisă sub forma

$$x^2 + ux + v = \sqrt{s}(rx + t),$$

unde r, s, t, u, v, sînt numere raţionale. Ridicînd la pătrat ambii membri, obţinem o ecuaţie de gradul patru

$$(x^2 + ux + v)^2 = s(rx + t)^2 \tag{5}$$

cu coeficienţi raţionali, aşa cum s-a afirmat.

Exerciţii: 1) Găsiţi ecuaţiile cu coeficienţi raţionali pentru a) $x = \sqrt{2 + \sqrt{3}}$; b) $x = \sqrt{2} + \sqrt{3}$; c) $x = 1/\sqrt{5 + \sqrt{3}}$.

2) Găsiţi printr-o metodă similară ecuaţiile de gradul opt pentru a) $x = \sqrt{2 + \sqrt{2 + \sqrt{2}}}$; b) $x = \sqrt{2} + \sqrt{1 + \sqrt{3}}$; c) $x = 1 + \sqrt{5 + \sqrt{3 + \sqrt{2}}}$.

Pentru a demonstra teorema în general, pentru un x cuprins într-un cîmp F_k, cu un k arbitrar, arătăm prin procedeul folosit mai sus că x satisface o ecuaţie pătratică, cu coeficienţi într-un cîmp F_{k-1}. Repetînd procedeul, găsim că x satisface o ecuaţie de gradul $2^2 = 4$ cu coeficienţi într-un cîmp F_{k-2} etc.

Exerciţiu: Completaţi demonstraţia generală, folosind inducţia matematică, pentru a arăta că x satisface o ecuaţie de gradul 2^l, cu coeficienţi într-un cîmp F_{k-l}, $0 < l \leqslant k$. Această afirmaţie, pentru $l = k$, este chiar teorema dorită.

*§ 3. NEREZOLVABILITATEA CELOR TREI PROBLEME GRECEŞTI

1. Dublarea cubului

Acum sîntem bine pregătiţi pentru studierea vechilor probleme ale trisecţiunii unghiului, dublării cubului şi construirii heptagonului regulat. Să considerăm mai întîi problema dublării cubului. Dacă cubul dat are o muchie de lungime unitate, volumul său va fi egal cu unitatea de volum; se cere să găsim muchia x a unui cub de volum dublu. De aceea, muchia x va satisface ecuaţia cubică simplă

$$x^3 - 2 = 0. \tag{1}$$

Demonstraţia noastră a faptului că acest număr nu poate fi construit doar cu rigla şi compasul este indirectă. Vom încerca să presupunem că este posibilă o construcţie. Conform discuţiei precedente, aceasta înseamnă că x se află într-un cîmp F_k, obţinut ca mai sus din cîmpul raţional prin extinderi succesive, efectuate prin ,,adjuncţionarea'' unor rădăcini pătrate. După cum vom arăta, această ipoteză duce la o consecinţă absurdă.

Ştim deja că x nu se poate afla în cîmpul raţional F_0, pentru că $\sqrt[3]{2}$ este un număr iraţional (cf. exerciţiul 1, p. 76). Deci x se poate afla numai într-o extindere F_k, unde k este un întreg pozitiv. Tot atît de bine putem presupune că k este *cel mai mic* întreg pozitiv, astfel încît x se află într-un F_k. Rezultă că x poate fi scris sub forma

$$x = p + q\sqrt{w},$$

unde p, q şi w aparţin unui F_{k-1}, dar $\sqrt{w}$ nu-i aparţine. Acum, printr-un raţionament algebric simplu dar important, vom arăta că dacă $x = p + q\sqrt{w}$ este o soluţie a ecuaţiei cubice (1), atunci $y = p - q\sqrt{w}$ este de asemenea o soluţie. Deoarece x se află în cîmpul F_k, x^3 şi $x^3 - 2$ se află de asemenea în F_k şi avem

$$x^3 - 2 = a + b\sqrt{w}, \tag{2}$$

unde a şi b se află în F_{k-1}. Printr-un calcul simplu putem arăta că $a = p^3 + 3pq^2w - 2$, $b = 3p^2q + q^3w$. Dacă punem

$$y = p - q\sqrt{w},$$

atunci înlocuind pe q prin $-q$ în aceste expresii ale lui a şi b, obţinem

$$y^3 - 2 = a - b\sqrt{w}. \tag{2'}$$

Dar x a fost prin ipoteză, o rădăcină a ecuației $x^3 - 2 = 0$, deci

$$(3) \qquad a + b\sqrt{w} = 0 .$$

Aceasta implică — și aici este cheia raționamentului — că a și b trebuie să fie ambele nule. Dacă b nu ar fi nul, am deduce din (3) că $\sqrt{w} = -a/b$. Dar atunci $\sqrt{w}$ ar fi un număr al cîmpului F_{k-1}, în care se află a și b, contrar ipotezei noastre. Deci $b = 0$, și din (3) rezultă imediat că și $a = 0$.

Acum, după ce am arătat că $a = b = 0$, din (2′) deducem imediat că $y = p - q\sqrt{w}$ este de asemenea o soluție a ecuației cubice (1), deoarece $y^3 - 2 = 0$. Mai mult $y \neq x$, adică $x - y \neq 0$, pentru că $x - y = 2q\sqrt{w}$ se poate anula numai dacă $q = 0$, și dacă ar fi așa, atunci $x = p$ s-ar afla în F_{k-1}, contrar ipotezei noastre.

Am arătat deci că dacă $x = p + q\sqrt{w}$ este o rădăcină a ecuației cubice (1), atunci $y = p - q\sqrt{w}$ este o rădăcină diferită a aceleiași ecuații. Aceasta duce imediat la o contradicție, deoarece există un singur număr real x care este rădăcină cubică a lui 2, celelalte rădăcini cubice ale lui 2 fiind imaginare (cf. p. 114); $y = p - q\sqrt{w}$ este în mod evident real, deoarece p, q și $\sqrt{w}$ erau reale.

Astfel, ipoteza noastră inițială a dus la o absurditate și deci am dovedit că este falsă; o soluție a ecuației (1) nu se poate afla într-un cîmp F_k, astfel încît dublarea cubului cu ajutorul riglei și compasului este imposibilă.

2. O teoremă asupra ecuațiilor cubice

Raționamentul algebric final era adaptat în mod special problemei particulare de care ne ocupăm. Dacă vrem să tratăm celelalte două probleme grecești, este de dorit să ne bazăm pe o teoremă mai generală. Din punct de vedere algebric, toate cele trei probleme depind de rezolvarea ecuațiilor cubice. Pe baza unui rezultat fundamental referitor la ecuația cubică

$$(4) \qquad z^3 + az^2 + bz + c = 0,$$

dacă x_1, x_2, x_3 sînt cele trei rădăcini ale acestei ecuații, atunci

$$(5) \qquad x_1 + x_2 + x_3 = -a^*.$$

---

* Polinomul $z^3 + az^2 + bz + c$ poate fi descompus în factori sub forma produsului $(z - x_1)(z - x_2)(z - x_3)$, unde x_1, x_2, x_3 sînt rădăcinile ecuației (4) (cf. p. 119). Deci

$$z^3 + az^2 + bz + c = z^3 - (x_1 + x_2 + x_3)z^2 + (x_1x_2 + x_1x_3 + x_2x_3)z - x_1x_2x_3,$$

astfel, deoarece coeficientul fiecărei puteri a lui z trebuie să fie același în ambii membri, deducem

$$-a = x_1 + x_2 + x_3, \quad b = x_1x_2 + x_1x_3 + x_2x_3, \quad c = -x_1x_2x_3.$$

Să considerăm o ecuație cubică (4) în care coeficienții a, b, c sînt numere raționale. Se poate ca una din rădăcinile ecuației să fie rațională; de exemplu, ecuația $x^3 - 1 = 0$ are rădăcina rațională 1, în timp ce celelalte două rădăcini date de ecuația pătratică $x^2 + x + 1 = 0$ sînt imaginare. Putem demonstra însă cu ușurință teorema generală: *Dacă o ecuație cubică cu coeficienți raționali nu are rădăcini raționale, atunci nici una din rădăcinile ei nu poate fi construită plecînd de la cîmpul rațional* F_0.

Demonstrația va fi dată ca și mai înainte printr-o metodă indirectă. Să presupunem că x este o rădăcină construibilă a ecuației (4). Atunci x s-ar afla în ultimul cîmp F_k al unui șir de extinderi $F_0, F_1, \ldots, F_k$ ca mai sus. Putem presupune că numărul k este *cel mai mic* întreg, astfel încît o rădăcină a ecuației cubice (4) se află într-o extindere F_k. Desigur, k trebuie să fie mai mare decît zero, deoarece în enunțul teoremei se presupune că nici o rădăcină x nu se află în cîmpul rațional F_0. Deci x poate fi scris sub forma

$$x = p + q\sqrt{w},$$

unde p, q, w se află în cîmpul precedent F_{k-1}, dar $\sqrt{w}$ nu se află în acest cîmp. Rezultă, ca și pentru ecuația particulară $z^3 - 2 = 0$ din secțiunea precedentă, că un alt număr din F_k, și anume

$$y = p - q\sqrt{w},$$

va fi de asemenea o rădăcină a ecuației (4). Ca și mai înainte, vedem că $q \neq 0$ și deci $x \neq y$.

Din (5) rezultă că a treia soluție u a ecuației (4) este dată de $u = -a - x - y$. Dar, deoarece $x + y = 2p$, aceasta înseamnă că

$$u = -a - 2p,$$

din care $\sqrt{w}$ a dispărut, astfel încît u este un număr din cîmpul F_{k-1}. Acest fapt contrazice ipoteza că numărul k este *cel mai mic număr*, astfel încît un F_k conține o rădăcină a ecuației (4). Prin urmare, ipoteza este absurdă și nici o rădăcină a ecuației (4) nu se poate afla într-un astfel de cîmp F_k. Teorema generală este demonstrată. Pe baza acestei teoreme se dovedește că este imposibilă construirea unei soluții doar cu ajutorul riglei și al compasului, dacă echivalentul algebric al problemei este rezolvarea unei ecuații cubice, care nu are rădăcini raționale. Această echivalență a fost evidentă pentru problema dublării cubului iar acum va fi stabilită pentru celelalte două probleme grecești.

3. Trisecțiunea unghiului

Vom demonstra că trisecțiunea unghiului, numai cu ajutorul riglei și compasului, este imposibilă *în general*. Desigur, există unghiuri, ca de pildă cele de 90 și de 180°, pentru care trisecțiunea poate fi efectuată. Ceea ce trebuie să

demonstrăm este faptul că trisecțiunea nu poate fi efectuată printr-un procedeu valabil pentru *orice* unghi. Pentru demonstrație este suficient să indicăm un singur unghi care nu poate fi trisecționat, deoarece o *metodă general valabilă* ar trebui să fie aplicabilă fiecărui caz particular. Prin urmare, inexistența unei metode generale va fi demonstrată dacă putem demonstra, de exemplu, că unghiul de 60° nu poate fi trisecționat numai cu ajutorul riglei și al compasului.

Putem obține un echivalent algebric al acestei probleme în diferite moduri; cel mai simplu este să presupunem că unghiul θ este dat prin cosinusul său: $\cos\theta = g$. Atunci problema este echivalentă cu aceea a găsirii cantității $x = \cos(\theta/3)$. Dintr-o formulă trigonometrică simplă (cf. p. 113), cosinusul lui $\theta/3$ este legat de cosinusul lui θ prin egalitatea

$$\cos\theta = g = 4\cos^3\frac{\theta}{3} - 3\cos\frac{\theta}{3}.$$

Cu alte cuvinte, problema trisecțiunii unghiului θ cu $\cos\theta = g$, se reduce la construirea unei soluții a ecuației cubice

(6) $$4z^3 - 3z - g = 0.$$

Pentru a arăta că acest lucru nu poate fi făcut în general, luăm $\theta = 60°$, astfel încît $g = \cos 60° = 1/2$. Ecuația (6) devine atunci

(7) $$8z^3 - 6z = 1.$$

În virtutea teoremei demonstrate în secțiunea precedentă, trebuie să arătăm doar că această ecuație nu are rădăcini raționale. Fie $v = 2z$. Atunci ecuația devine

(8) $$v^3 - 3v = 1.$$

Dacă ar exista un număr rațional $v = r/s$, care satisface această ecuație, unde r și s sînt întregi relativ primi, ar trebui să avem $r^3 - 3s^2r = s^3$. De aici rezultă că $s^3 = r(r^2 - 3s^2)$ se divide cu r, ceea ce înseamnă că r și s au un factor comun, cu excepția cazului în care $r = \pm 1$. De asemenea, s^2 este un divizor al lui $r^3 = s^2(s + 3r)$, ceea ce înseamnă că r și s au un divizor comun, cu excepția cazului în care $s = \pm 1$. Deoarece am presupus că r și s nu au divizori comuni, am arătat că singurele numere raționale care ar putea satisface ecuația (8) sînt $+1$ sau -1. Substituind pe $+1$ și pe -1 în locul lui v în ecuația (8), vedem că nici una din aceste valori nu o satisface. Deci ecuația (8), și prin urmare ecuația (7), nu au rădăcini raționale, ceea ce demonstrează imposibilitatea trisecțiunii unghiului.

Teorema că unghiul în general nu poate fi trisecționat cu ajutorul riglei și al compasului este adevărată, numai dacă rigla este privită doar ca instrument de trasare a unei drepte prin două puncte date. În caracterizarea generală a numerelor construibile pe care am prezentat-o, fo-

losirea riglei era limitată întotdeauna doar la această operaţie. Permiţînd alte utilizări ale riglei, totalitatea construcţiilor posibile poate fi extinsă într-o măsură considerabilă. Metoda următoare pentru trisecţionarea unghiului, găsită în lucrările lui Arhimede, constituie un bun exemplu.

Fie dat un unghi arbitrar x, ca în fig. 36. Să prelungim una din laturile unghiului spre stînga şi să trasăm un semicerc cu centrul în O şi de rază arbitrară r. Să determinăm două puncte A şi B pe muchia riglei, astfel încît $AB = r$. Păstrînd punctul B pe semicerc, să glisăm rigla în

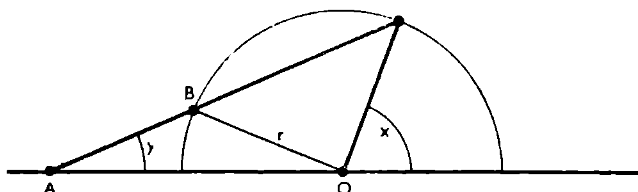


Fig. 36. Trisecţiunea unui unghi propusă de Arhimede

poziţia în care A se află pe latura prelungită a unghiului x, în timp ce muchia riglei trece prin intersecţia celei de-a doua laturi a unghiului x, cu semicercul trasat din O. Cu rigla în această poziţie, să trasăm o dreaptă, care face un unghi y cu latura prelungită a unghiului iniţial x.

Exerciţiu: Arătaţi că această construcţie ne dă într-adevăr $y = x/3$.

4. Heptagonul regulat

Să considerăm acum problema găsirii laturii x a unui heptagon regulat, înscris în cercul unitate. Cel mai simplu mod de rezolvare a acestei probleme foloseşte numerele complexe (cf. cap. II, § 5). Ştim că vîrfurile heptagonului sînt date de rădăcinile ecuaţiei

$$(9) \qquad x^7 - 1 = 0,$$

coordonatele x, y ale vîrfurilor fiind considerate ca părţi reale şi imaginare ale numerelor complexe $z = x + yi$. O rădăcină a acestei ecuaţii este $z = 1$, iar celelalte sînt rădăcinile ecuaţiei

$$(10) \qquad \frac{z^7 - 1}{z - 1} = z^6 + z^5 + z^4 + z^3 + z^2 + z + 1 = 0,$$

obţinută din (9) prin simplificare cu factorul $z - 1$ (cf. p. 116). Împărţind ecuaţia (10) prin z^3, obţinem ecuaţia

$$(11) \qquad z^3 + 1/z^3 + z^2 + 1/z^2 + z + 1/z + 1 = 0.$$

Printr-o transformare algebrică simplă, aceasta poate fi scrisă sub forma

(12) $$(z + 1/z)^3 - 3(z + 1/z) + (z + 1/z)^2 - 2 + (z + 1/z) + 1 = 0 .$$

Notînd cantitatea $z + 1/z$ cu y, găsim din (12) că

(13) $$y^3 + y^2 - 2y - 1 = 0.$$

Știm că z, rădăcina de ordinul șapte a unității, este dată de

(14) $$z = \cos\varphi + i \sin\varphi,$$

unde $\varphi = 360°/7$ este unghiul subîntins în centrul cercului de latura heptagonului regulat; de asemenea, din exercițiul 2, p. 114, știm că $1/z = \cos\varphi - i \sin\varphi$, astfel încît $y = z + 1/z = 2 \cos\varphi$. Dacă putem construi pe y, atunci putem construi și pe $\cos\varphi$, și reciproc. Deci, dacă putem demonstra că y nu este construibil, vom arăta în același timp că z, și deci și heptagonul, nu este construibil. Astfel, considerînd teorema din secțiunea 2, rămîne să arătăm doar că ecuația (13) nu are rădăcini raționale, și acest lucru se demonstrează în mod indirect. Să admitem că ecuația (13) are o rădăcină rațională r/s, unde r și s sînt întregi care nu au nici un factor comun. Atunci avem

(15) $$r^3 + r^2s - 2rs^2 - s^3 = 0;$$

de unde, ca și mai sus, vedem că r^3 are factorul s, iar s^3 are factorul r. Deoarece r și s nu au nici un factor comun, fiecare trebuie să fie egal cu ± 1; de aceea, y poate avea numai valorile $+1$ și -1, dacă este rațional. Substituind aceste numere în ecuație, vedem că nici unul din ele nu o satisface. Deci y și, prin urmare, latura heptagonului regulat, nu sînt construibile.

5. Observații asupra problemei cvadraturii cercului

Am reușit să rezolvăm problemele dublării cubului, trisecțiunii unghiului și construcției heptagonului regulat prin metode destul de elementare. Problema cvadraturii cercului este mult mai dificilă și necesită tehnica analizei matematice avansate. Deoarece un cerc de rază r are aria egală cu πr^2, problema construirii unui pătrat cu o arie egală cu aceea a unui cerc dat, a cărui rază este egală cu unitatea de lungime, se reduce la construirea unui segment de lungime $\sqrt{\pi}$, ca latură a pătratului căutat. Acest segment va fi construibil dacă și numai dacă numărul π este construibil. În lumina caracterizării noastre generale a numerelor construibile, am putea demonstra imposibilitatea cvadraturii cercului, arătînd că numărul π nu poate fi conținut în nici un cîmp F_k, care poate fi obținut prin „adjuncționări" succesive ale rădăcinilor pătrate la cîmpul rațional F_0. Deoarece toate elementele unui astfel de cîmp sînt numere algebrice, adică numere care satisfac ecuații algebrice cu coefi-

cienți întregi, va fi suficient să arătăm că numărul π nu este algebric, adică este transcendent (cf. p. 120).

Aparatul tehnic necesar pentru demonstrarea faptului că π este un număr transcendent a fost creat de Charles Hermite (1822—1905), care a demonstrat că numărul e este transcendent. Printr-o ușoară extindere a metodei lui Hermite, F. Lindemann a reușit (1882) să demonstreze transcendența lui π, rezolvînd astfel definitiv problema milenară a cvadraturii cercului. Demonstrația lui Lindemann este accesibilă celor care au studiat analiza matematică, dar depășește cadrul acestei cărți.

PARTEA A II-A

DIFERITE METODE DE EFECTUARE A CONSTRUCȚIILOR

§ 4. TRANSFORMĂRI GEOMETRICE. INVERSIUNEA

1. Observații generale

În cea de-a doua parte a acestui capitol, vom discuta sistematic cîteva principii generale, care pot fi aplicate problemelor de construcție. Multe dintre aceste probleme pot fi privite mai clar, din punctul de vedere general al „transformărilor geometrice"; în loc de a studia o construcție particulară, vom considera simultan o întreagă clasă de probleme legate între ele prin anumite procese de transformare. Puterea clarificatoare a noțiunii de clasă de transformări geometrice nu este restrînsă cîtuși de puțin la problemele de construcție, ci influențează aproape întreaga geometrie. În capitolele IV și V vom trata acest aspect general al transformărilor geometrice. Aici vom studia un tip particular de transformare, inversiunea în raport cu un cerc dintr-un plan, care este o generalizare a simetriei obișnuite față de o dreaptă.

Prin *transformare* sau *aplicație* a planului pe el însuși, înțelegem o regulă care asociază fiecărui punct P al planului un alt punct P', numit *imaginea* lui P prin transformare; punctul P se numește *antecedentul*[1] lui P'. Un exemplu simplu de astfel de transformare este dat de *simetria* față de o dreaptă L, aflată în plan: un punct P aflat de o parte a lui L are ca imagine punctul P', aflat de cealaltă parte a lui L, astfel încît L este mediatoarea segmentului PP'. O transformare poate lăsa pe loc anumite puncte ale planului; în cazul simetriei, acest lucru se întîmplă cu punctele de pe dreapta L.

[1] Cu acest sens, termenul nu este încetățenit în limba română. — *N.T.*

Alte exemple de transformări sînt *rotațiile* planului în jurul unui punct fix O, *translațiile* paralele, care deplasează fiecare punct cu o distanță d, într-o direcție și într-un sens dat (o astfel de transformare nu are puncte fixe) și, mai general, *mișcările rigide* ale planului, care pot fi gîndite ca fiind compuse din rotații și translații paralele.

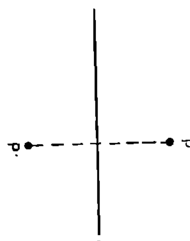


Fig. 37. Simetricul unui punct față de o dreaptă

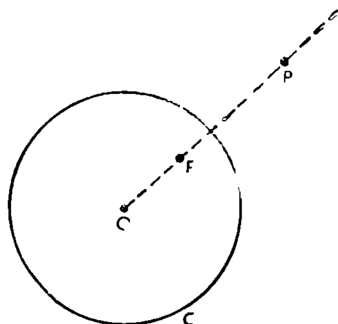


Fig. 38. Inversul unui punct față de un cerc

Clasa de transformări particulare care ne interesează acum este formată de *inversiunile* în raport cu cercuri. (Acestea sînt cunoscute uneori sub numele de simetrii circulare, pentru că cu o oarecare aproximație, ele reprezintă relația dintre original și imagine în reflexia într-o oglindă circulară). Într-un plan fix, fie dat un cerc C, cu centrul O (numit polul inversiunii) și de rază r. Imaginea unui punct P este, prin definiție, punctul P' care se află pe semidreapta OP, astfel încît

$$OP \cdot OP' = r^2. \quad (1)$$

Se spune că punctele P și P' sînt *puncte inverse* în raport cu C. Din această definiție rezultă că dacă P' este inversul punctului P, atunci P este inversul lui P'. O inversiune permută interiorul și exteriorul cercului C, deoarece dacă $OP < r$, avem $OP' > r$, iar pentru $OP > r$, avem $OP' < r$, Singurele puncte ale planului, care rămîn fixe prin transformare, sînt punctele de pe cercul C însuși.

Regula (1) nu definește o imagine a centrului O. Este clar că dacă un punct variabil P se apropie de O, imaginea P' se va îndepărta din ce în ce mai mult. Din acest motiv, spunem uneori că punctul O însuși corespunde prin inversiune *punctului de la infinit* Utilitatea acestei terminologii se află în faptul că ea ne permite să afirmăm că o inversiune stabilește o corespondență între punctele

planului şi imaginile lor, care este biunivocă fără excepţie: orice punct al planului are o singură imagine şi este el însuşi imaginea unui singur punct. Această proprietate este împărtăşită de toate transformările considerate mai înainte.

2. Proprietăţile inversiunii

Cea mai importantă proprietate a unei inversiuni este faptul că ea transformă dreptele şi cercurile în drepte şi cercuri. Mai precis, vom arăta că printr-o inversiune

(a) o dreaptă care trece prin O devine o dreaptă care trece prin O
(b) o dreaptă care nu trece prin O devine un cerc care trece prin O
(c) un cerc care trece prin O devine o dreaptă care nu trece prin O
(d) un cerc care nu trece prin O devine un cerc care nu trece prin O.

Propoziţia (a) este evidentă, deoarece din definiţia inversiunii, orice punct de pe dreaptă are ca imagine un alt punct al aceleiaşi drepte, astfel încît, cu toate că punctele dreptei sînt permutate, dreapta în ansamblu se transformă în ea însăşi.

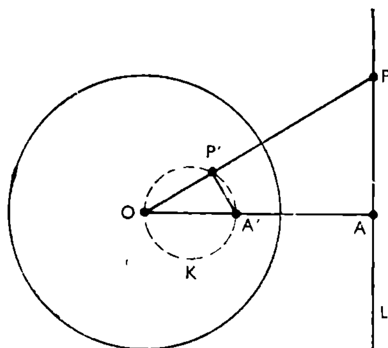


Fig. 39. Transformarea prin inversiune a unei drepte L *într-un cerc*

Pentru a demonstra propoziţia (b), să coborîm o perpendiculară din O pe dreapta L (fig. 39.) Fie A piciorul acestei perpendiculare şi fie A' inversul punctului A. Fie P un punct oarecare de pe L şi fie P' inversul său. Deoarece $OA' \cdot OA = OP' \cdot OP = r^2$, rezultă că

$$\frac{OA'}{OP'} = \frac{OP}{OA}.$$

Deci, triunghiurile $OP'A'$ şi OAP sînt asemenea şi unghiul $OP'A'$ este drept. Din geometria elementară rezultă că P' se află pe cercul K, de diametru OA', astfel încît figura inversă a lui L este acest cerc. Acest fapt demonstrează propoziţia (b). Propoziţia (c) rezultă acum din faptul că, deoarece inversul lui L este K, inversul lui K este L.

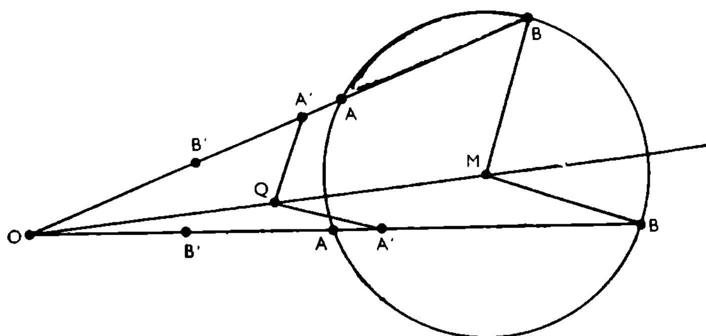


Fig. 40. Inversiunea unui cerc

Rămîne să demonstrăm propoziţia (d). Fie K un cerc care nu trece prin O, cu centrul în M şi de rază k. Pentru a obţine imaginea lui, ducem o dreaptă prin O, care intersectează pe K în A şi B, şi apoi stabilim cum variază imaginile A', B', cînd dreapta dusă prin O intersectează pe K în toate modurile posibile. Să notăm distanţele OA, OB, OA', OB', OM cu a, b, a', b', m şi fie t lungimea unei tangente dusă prin O la K. Avem $aa' = bb' = r^2$, în baza definiţiei inversiunii şi $ab = t^2$, în baza unei proprietăţi geometrice elementare a cercului. Dacă împărţim primele egalităţi prin ultima obţinem

$$a'/b = b'/a = r^2/t^2 = c^2,$$

unde c^2 este o constantă care depinde numai de r şi t, şi este aceeaşi pentru toate poziţiile lui A şi B. Prin A' ducem o dreaptă paralelă cu BM, care întîlneşte pe OM în Q. Fie $OQ = q$ şi $A'Q = \rho$. Atunci $q/m = a'/b = \rho/k$, sau

$$q = ma'/b = mc^2, \quad \rho = ka'/b = kc^2.$$

Aceasta înseamnă că pentru toate poziţiile lui A şi B, Q se va afla întotdeauna în acelaşi loc pe OM, şi distanţa $A'Q$ va avea întotdeauna aceeaşi valoare. De asemenea, $B'Q = \rho$, deoarece $a'/b = b'/a$. Astfel, imaginile tuturor punctelor A, B, de pe K sînt puncte a căror distanţă la Q este întotdeauna egală cu ρ, adică imaginea lui K este un cerc. Aceasta demonstrează propoziţia (d).

3. Construcția geometrică a punctelor inverse

Teorema următoare va fi utilă în secțiunea 4 a acestui paragraf: *Punctul* P′, *inversul unui punct dat* P *în raport cu un cerc* C, *poate fi construit geometric doar cu ajutorul compasului.* Să considerăm mai întîi cazul în care punctul dat P este exterior cercului C. Cu OP ca rază și P ca centru, să descriem un

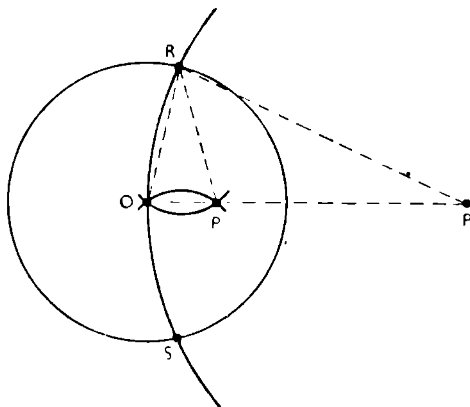


Fig. 41. Inversiunea unui punct exterior în raport cu un cerc

arc de cerc, care intersectează pe C în punctele R și S. Cu aceste două puncte luate ca centre, să descriem arce de cerc de rază r, care se intersectează în O și într-un punct P' de pe dreapta OP. În triunghiurile isoscele ORP și ORP' avem

$$\sphericalangle ORP = \sphericalangle POR = \sphericalangle OP'R,$$

astfel încît aceste triunghiuri sînt asemenea, și de aceea

$$\frac{OP}{OR} = \frac{OR}{OP'}, \text{ adică } OP \cdot OP' = r^2.$$

Deci P' este inversul căutat al lui P, care trebuia construit.

Dacă punctul dat P se află în interiorul lui C, aceeași construcție și demonstrație rămîn în vigoare, cu condiția ca cercul de rază OP, și cu centrul în P, să intersecteze pe C în două puncte. În caz contrar, putem reduce construirea punctului invers P' la cazul precedent, folosind următorul artificiu simplu.

Să observăm mai întîi că, doar cu ajutorul compasului, putem găsi un punct C pe dreapta care unește două puncte date A, O, astfel încît $AO = OC$. Pentru a face acest lucru, să trasăm un cerc cu centrul în O, de rază $r = AO$, și să determinăm pe acest cerc punctele P, Q, C, astfel încît $AP = PQ = QC = r$.

Atunci C este punctul dorit, după cum se vede din faptul că triunghiurile AOP, OPQ, OQC sînt echilaterale, astfel încît OA şi OC formează un unghi de $180°$ şi $OC = OQ = AO$. Repetînd acest procedeu, putem prelungi cu uşurință pe AO ori de cîte ori dorim. În treacăt fie zis, deoarece lungimea segmentului AO este egală cu $r\sqrt{3}$, după cum cititorul poate verifica cu uşurință, am construit în acelaşi timp pe $\sqrt{3}$, pornind de la segmentul unitate, fără a folosi rigla.

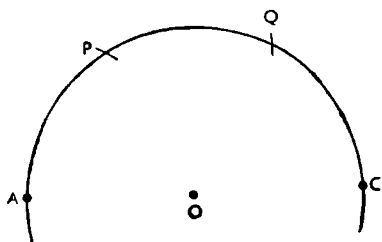


Fig. 42. Dublarea unui segment

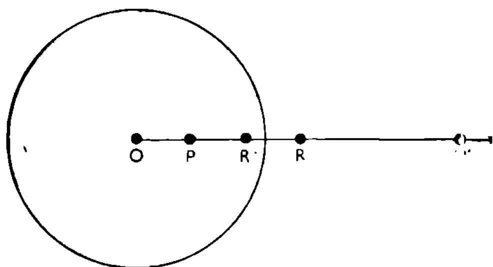


Fig. 43. Inversiunea unui punct interior în raport cu un cerc

Acum putem găsi inversul oricărui punct P aflat în interiorul cercului C. Mai întîi să găsim un punct R pe dreapta OP, a cărui distanță pînă la O este un multiplu întreg al lui OP şi care să se afle în exteriorul lui C,

$$OR = n \cdot OP.$$

Putem face acest lucru purtînd succesiv distanța OP, cu compasul, pînă ce ajungem în exteriorul lui C. Acum găsim punctul R', inversul lui R, prin construcția dată mai înainte. Atunci

$$r^2 = OR' \cdot OR = OR' \cdot (n \cdot OP) = (n \cdot OR') \cdot OP.$$

De aceea, punctul P', pentru care $OP' = n \cdot OR'$, este inversul dorit.

4. Cum putem împărți un segment în două părți egale şi găsi centrul unui cerc doar cu ajutorul compasului

Acum, după ce am aflat cum putem găsi inversul unui punct doar cu ajutorul compasului, putem efectua cîteva construcții interesante. De exemplu, să considerăm problema găsirii mijlocului unui segment cu extremitățile A şi B, folosind doar compasul (nu se poate trasa nici o dreaptă !). Iată soluția : să trasăm cercul de rază AB, cu centrul B, şi să determinăm trei arce de rază AB, începînd din A. Punctul final C se va afla pe dreapta AB, aşa că

$AB = BC$. Să trasăm acum cercul de rază AB şi de centru A, şi fie C' inversul punctului C în raport cu acest cerc. Atunci

$$AC' \cdot AC = AB^2, \; AC' \cdot 2AB = AB^2, \; 2AC' = AB.$$

Deci C' este mijlocul dorit.

O altă construcţie cu ajutorul compasului, care foloseşte puncte inverse, este aceea prin care găsim centrul unui cerc, în care cunoaştem doar circum-

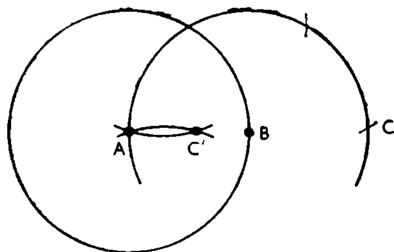


Fig. 44. Aflarea mijlocului unui segment

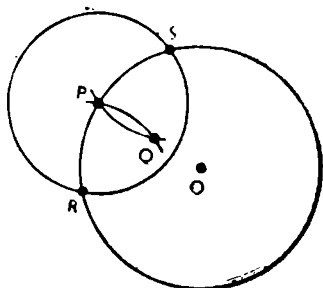


Fig. 45. Aflarea centrului unui cerc

ferinţa. Alegem un punct oarecare P pe circumferinţă şi, cu centrul în acest punct, trasăm un cerc care intersectează cercul dat în punctele R şi S. Cu centrele în aceste puncte, trasăm arce de raze $RP = SP$, care se intersectează în punctul Q. O comparaţie cu fig. 41 arată că centrul necunoscut Q' este inversul lui Q în raport cu cercul de centru P, astfel încît Q' poate fi construit doar cu ajutorul compasului.

§ 5. CONSTRUCŢII EFECTUATE CU AJUTORUL ALTOR INSTRUMENTE. CONSTRUCŢIILE LUI MASCHERONI EFECTUATE DOAR CU AJUTORUL COMPASULUI

*1. O construcţie clasică pentru dublarea cubului

Pînă acum, am considerat doar probleme de construcţie geometrică, care folosesc numai rigla şi compasul. Dacă sînt permise şi alte instrumente, atunci mulţimea construcţiilor posibile se măreşte. De exemplu, grecii au rezolvat problema dublării cubului în modul următor. Să considerăm (ca în fig. 46) un unghi drept rigid MZN şi o cruce rectangulară şi mobilă B, VW,PQ. Două muchii suplimentare RS şi TU pot aluneca perpendicular pe laturile unghiului drept. Să alegem pe cruce două puncte fixe E şi G, astfel încît $GB = a$ şi $BE = f$ să aibă lungimi date dinainte. Aşezînd crucea în aşa fel, încît punctele E şi G să se afle respectiv pe NZ şi MZ şi glisînd muchiile TU şi RS, pu-

tem aduce întregul aparat într-o poziţie în care avem un dreptunghi $ADEZ$ prin ale cărui vîrfuri A, D, E trec braţele BW, BQ, BV ale crucii. O astfel de aşezare este întotdeauna posibilă dacă $f > a$. Vedem imediat că $a : x = x : y = y : f$, deci dacă punem în aparat $f = 2a$, avem $x^3 = 2a^3$. Prin urmare, x va fi muchia unui cub, al cărui volum este dublul volumului cubului cu muchia a. Tocmai acest lucru se cerea la dublarea cubului.

2. Construcţii numai cu ajutorul compasului

În timp ce este natural ca acceptînd un număr mai mare de instrumente să putem rezolva un număr mai mare de probleme de construcţie, ne-am putea aştepta că reducînd numărul instrumentelor permise, să micşorăm clasa construcţiilor posibile. De aceea, a fost foarte surprinzătoare descoperirea făcută

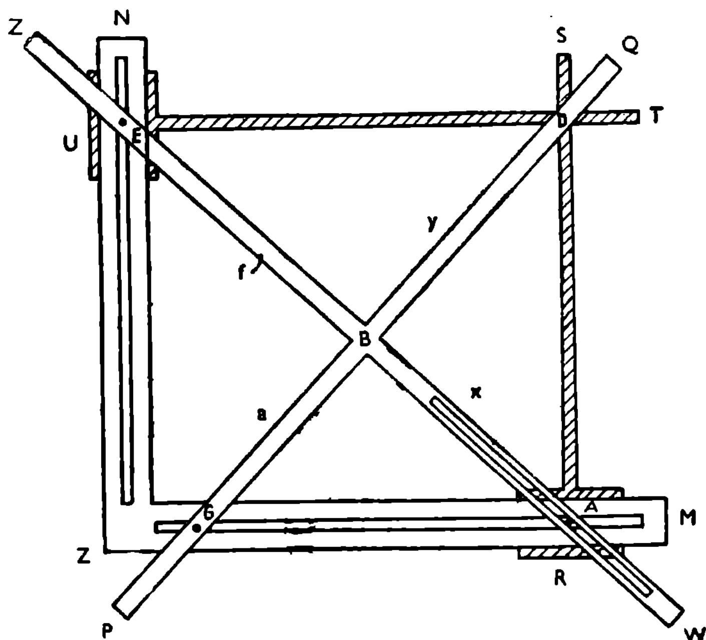


Fig. 46. Instrument pentru dublarea cubului

de italianul Mascheroni (1750—1800) că *toate construcţiile geometrice posibile doar cu ajutorul riglei şi al compasului pot fi efectuate numai cu compasul.* Desigur, nu putem trasa dreapta care uneşte două puncte, dacă nu dispunem de riglă, astfel încît această construcţie fundamentală nu este posibilă în teoria lui Mascheroni. În schimb, trebuie să concepem o dreaptă ca fiind dată de

oricare două din punctele ei. Folosind doar compasul, pe această cale putem găsi punctul de intersecție a două drepte și de asemenea intersecția unui cerc dat cu o dreaptă.

Poate cel mai simplu exemplu de construcție Mascheroni este dublarea unui segment dat AB. Soluția a fost dată la p. 163. La p. 164 am împărțit în două părți egale un segment. Acum vom rezolva problema înjumătățirii unui arc dat

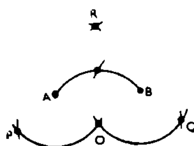


Fig. 47. Înjumătățirea unui arc cu ajutorul compasului

AB, al unui cerc de centru O. Construcția este următoarea: din A și B ca centre să trasăm două arce de cerc, de rază AO. Din O să ducem arcele OP și OQ, egale cu AB. Apoi să trasăm două arce cu PB și QA ca raze și cu P și Q ca centre, care se intersectează în R. În sfîrșit, cu OR ca rază, să descriem un arc cu centrul în P sau Q, pînă ce intersectăm pe AB. Acest punct de intersecție este mijlocul cerut al arcului AB. Demonstrația este lăsată pe seama cititorului, ca exercițiu.

Ar fi imposibil să demonstrăm teorema generală a lui Mascheroni dînd efectiv o construcție, doar cu ajutorul compasului, pentru fiecare construcție posibilă cu rigla și compasul, pentru că numărul construcțiilor posibile nu este finit. Însă putem ajunge la același scop, demonstrînd că fiecare dintre următoarele patru construcții fundamentale este posibilă doar cu ajutorul compasului.

1. Trasarea unui cerc de rază și de centru date.
2. Găsirea punctelor de intersecție a două cercuri.
3. Găsirea punctelor de intersecție ale unei drepte cu un cerc.
4. Găsirea punctelor de intersecție a două drepte. Orice construcție geometrică în sensul obișnuit, cu ajutorul riglei și al compasului, constă dintr-o succesiune finită a acestor construcții elementare. Primele două sînt evident posibile doar cu ajutorul compasului. Rezolvarea problemelor mai dificile 3 și 4 depinde de proprietățile inversiunii, dezvoltate în paragraful precedent.

Să rezolvăm problema 3: să găsim punctele de intersecție ale unui cerc C, cu o dreaptă dată prin două puncte A și B. Cu centrele A și B și respectiv razele AO și BO să trasăm două arce, care se intersectează din nou în P. Să determinăm acum punctul Q, inversul lui P în raport cu C, prin construcția dată la p. 164 și efectuată doar cu ajutorul compasului. Să trasăm cercul cu centrul Q și raza QO (acest cerc trebuie să intersecteze pe C); punctele de intersecție X și X' ale acestui cerc cu cercul C dat sînt punctele cerute. Pentru

a demonstra acest lucru trebuie să arătăm doar că X şi X' sînt echidistante faţă de O şi P, dat fiind că A şi B au această proprietate prin construcţie. Aceasta rezultă din faptul că inversul lui Q este un punct a cărui distanţă la X şi X' este egală cu raza lui C (p. 162). Să remarcăm că cercul care trece prin X, X' şi O este inversul dreptei AB, deoarece acest cerc şi dreapta AB intersectează cercul C în aceleaşi puncte (punctele de pe circumferinţa unui cerc sînt propriile lor inverse).

Construcţia este imposibilă numai în cazul cînd dreapta AB trece prin centrul cercului C. Dar atunci punctele de intersecţie pot fi găsite prin construcţia dată la p. 166, ca mijloace ale arcelor de pe C, obţinute prin trasarea unui cerc arbitrar cu centrul în B, care intersectează cercul C în B_1 şi B_2.

Metoda determinării cercului invers dreptei care uneşte două puncte date permite o rezolvare imediată a problemei 4. Să presupunem că dreptele sînt date prin AB şi $A'B'$ (fig. 50). Să trasăm un cerc oarecare C în plan, şi prin metoda precedentă să găsim cercurile inverse dreptelor AB şi $A'B'$. Aceste cercuri se intersectează în O şi într-un punct Y. Punctul X, inversul lui Y, este punctul de intersecţie cerut şi poate fi construit prin procedeul deja folosit. Faptul că X este punctul cerut este evident, deoarece Y este singurul punct invers, atît al unui punct al lui AB, cît şi al unui punct de pe $A'B'$: deci punctul X, inversul lui Y, trebuie să se afle atît pe AB, cît şi pe $A'B'$.

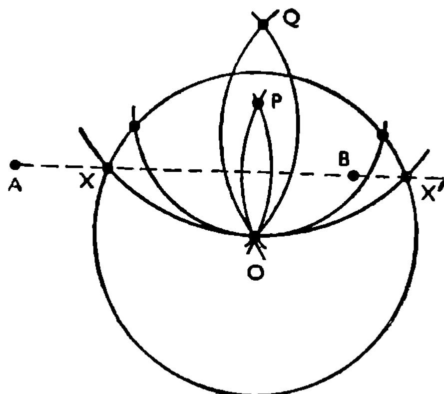


Fig. 48. Intersecţia unui cerc cu o dreaptă care nu trece prin centrul cercului

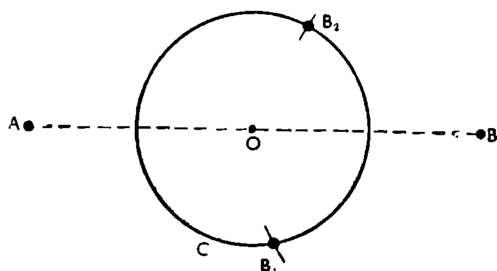


Fig. 49. Intersecţia unui cerc cu o dreaptă care trece prin centrul cercului

Cu aceste două construcţii am completat demonstraţia echivalenţei dintre construcţiile lui Mascheroni, care utilizează numai compasul, şi construcţiile geometrice clasice cu rigla şi compasul. Nu ne-am străduit să găsim soluţii elegante pentru fiecare problemă în parte, deoarece scopul nostru a fost mai

degrabă să furnizăm o privire în cadrul general al construcțiilor lui Mascheroni. Vom da însă ca exemplu construcția pentagonului regulat. Mai precis, vom găsi cinci puncte pe un cerc, care vor fi vîrfurile unui pentagon regulat înscris.

Fie A un punct oarecare al cercului dat K. Latura unui hexagon regulat înscris este egală cu raza lui K. Prin urmare, putem găsi punctele B, C, D pe K, astfel încît $\widehat{AB} = \widehat{BC} = \widehat{CD} = 60°$ (fig. 51). Cu A și D ca centre și AC ca rază trasăm arce care se intersectează pe X. Atunci, dacă O este centrul lui K, un arc cu centrul în A, de rază OX, va intersecta pe K în mijlocul F al arcului $\widehat{BC}$ (cf. p. 166). Acum, cu raza lui K, trasăm arce de centru F, care intersectează cercul K în G și H. Fie Y un punct a cărui distanță pînă la G și H este egală cu OX, astfel încît O să se afle între X și Y. Atunci segmentul AY va fi egal cu latura pentagonului cerut. Demonstrația este lăsată ca exercițiu pe seama cititorului. Remarcați că în construcție au fost folosite doar trei raze diferite.

În 1928, matematicianul danez Hjelmslev a găsit în Copenhaga un exemplar al cărții *Euclides Danicus*, publicată în 1672 de un autor obscur, G. Mohr. Din titlu s-ar putea deduce că această lucrare ar fi doar o versiune sau un

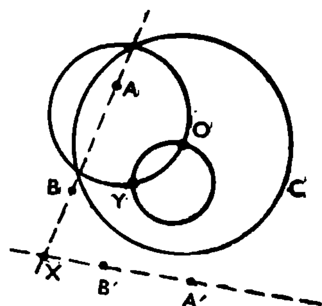


Fig. 50. Intersecția a două drepte

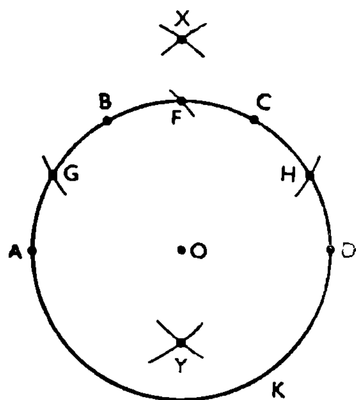


Fig. 51. Construirea pentagonului regulat

comentariu la *Elementele* lui Euclid. Dar atunci cînd Hjelmslev examină cartea, spre surprinderea sa, a găsit că ea conținea în esență problema lui Mascheroni și rezolvarea ei completă, găsită cu mult înaintea lui Mascheroni.

Exerciții: În cele ce urmează descriem construcțiile lui Mohr. Verificați valabilitatea lor. De ce rezolvă ele problemele lui Mascheroni?

1) Pe un segment AB de lungime p ridicaţi un segment perpendicular BC. (Indicaţie: prelungiţi segmentul AB printr-un punct D, astfel încît $AB = BD$. Trasaţi cercuri arbitrare cu centrele în A şi D, determinînd în acest mod pe C.)

2) Sînt date, undeva în plan, două segmente de lungimi p şi q, cu $p > q$. Găsiţi un segment de lungime $x = \sqrt{p^2 - q^2}$, folosind rezultatul precedent.

3) Dintr-un segment dat a, construiţi segmentul $a\sqrt{2}$. (Indicaţie: observaţi că $(a\sqrt{2})^2 = (a\sqrt{3})^2 - a^2$.)

4) Date fiind segmentele p şi q, găsiţi un segment $x = \sqrt{p^2 + q^2}$. (Indicaţie: folosiţi relaţia $x^2 = 2p^2 - (p^2 - q^2)$.) Găsiţi alte construcţii similare.

5) Folosind rezultatele precedente, găsiţi segmente de lungime $p + q$ şi $p - q$, date fiind segmentele de lungime p şi q, situate undeva în plan.

6) Verificaţi şi demonstraţi următoarea construcţie pentru mijlocul M al unui segment AB' de lungime a. Pe prelungirea lui AB găsiţi pe C şi D, astfel încît $CA = AB = BD$. Construiţi triunghiul isoscel ECD cu $EC = ED = 2a$ şi găsiţi pe M ca intersecţie a cercurilor cu diametrii EC şi ED.

7) Găsiţi proiecţia ortogonală a unui punct A pe o dreaptă BC.

8) Găsiţi pe x, astfel încît $x : a = p : q$, a, p şi q fiind segmente date.

9) Găsiţi pe $x = ab$, a şi b fiind segmente date.

Inspirat de Mascheroni, Jacob Steiner (1796—1863) a încercat să folosească doar rigla în locul compasului. Desigur, doar cu ajutorul riglei nu putem ieşi dintr-un cîmp numeric dat şi deci ea nu este suficientă pentru toate construcţiile geometrice în sensul clasic. Este cu atît mai remarcabil faptul că Steiner a putut reduce folosirea compasului la o singură aplicare. El a demonstrat că toate construcţiile din plan, care sînt posibile doar cu rigla şi compasul, sînt posibile doar cu ajutorul riglei, cu condiţia ca să fie dat un singur cerc fixat şi centrul său. Aceste construcţii utilizează metode proiective şi vor fi indicate mai tîrziu (cf. p. 216).

Nu putem renunţa la acest cerc şi la centrul său. De exemplu, dacă este dat un cerc, dar nu şi centrul său, acesta nu poate fi construit doar cu ajutorul riglei. Pentru a demonstra acest lucru, vom folosi un fapt care va fi tratat mai tîrziu (p. 238): există o transformare a planului în el însuşi, care are următoarele proprietăţi: (a) cercul dat este invariant prin transformare, (b) orice dreaptă este transformată într-o dreaptă, (c) centrul cercului este transformat într-un alt punct. Simpla existenţă a unei astfel de transformări demonstrează imposibilitatea construirii doar cu ajutorul riglei a centrului cercului dat. Pentru că oricare ar fi construcţia, ea ar consta din trasarea unui anumit număr de drepte şi din găsirea intersecţiilor lor, cu ele însele sau cu cercul dat. Acum, dacă întreaga figură formată din cercul dat, împreună cu punctele şi dreptele construcţiei, este supusă transformării a cărei existenţă am presupus-o, figura transformată va satisface toate cerinţele construcţiei, dar va da ca rezultat un punct diferit de centrul cercului dat. Deci o astfel de construcţie este imposibilă.

3. Desenarea cu ajutorul instrumentelor mecanice. Curbe mecanice. Cicloide

Inventînd mecanisme pentru trasarea curbelor, diferite de cerc și de dreaptă, putem lărgi în mare măsură domeniul figurilor construibile. De exemplu, dacă avem un instrument pentru desenarea hiperbolelor $xy = k$ și un altul pentru desenarea parabolelor $y = ax^2 + bx + c$, atunci orice problemă care duce la o ecuație cubică

$$ax^3 + bx^2 + cx = k, \tag{1}$$

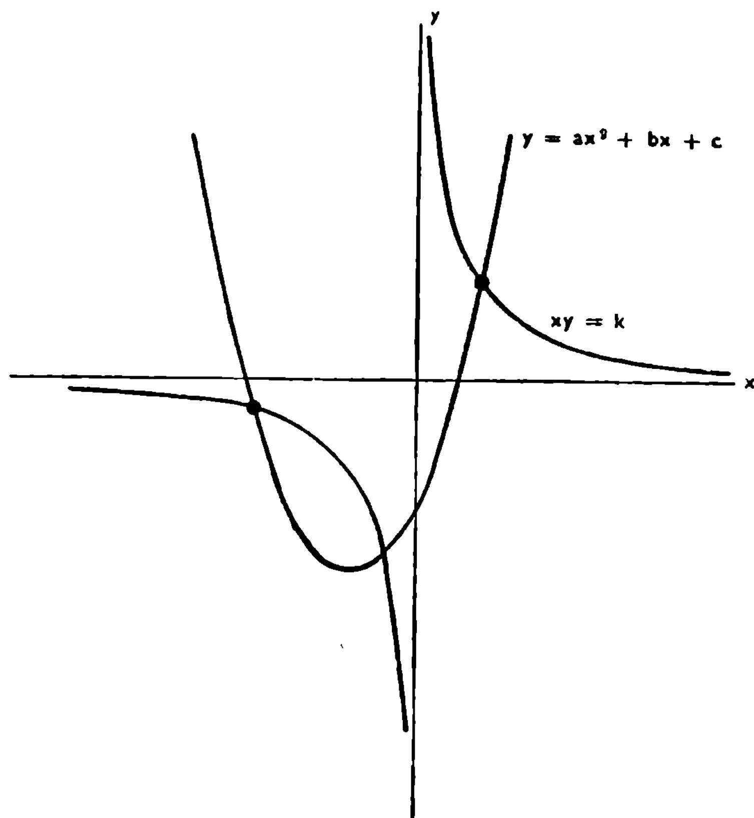


Fig. 52. Rezolvarea grafică a unei ecuații cubice

poate fi rezolvată printr-o construcție care folosește doar aceste instrumente, deoarece dacă notăm

$$xy = k, \quad y = ax^2 + bx + c, \tag{2}$$

atunci rezolvarea ecuației (1) se reduce la rezolvarea sistemului (2) prin eliminarea lui y; adică rădăcinile ecuației (1) sînt abscisele punctelor de intersecție ale hiperbolei și parabolei din (2). Astfel, soluțiile ecuației (1) pot fi

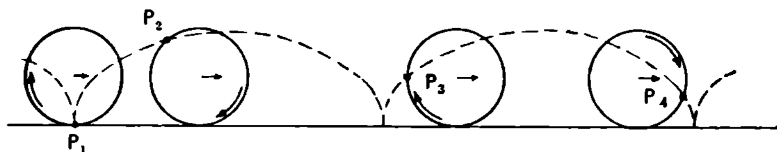


Fig. 53. Cicloida

construite, dacă avem instrumente cu ajutorul cărora putem trasa hiperbola și parabola, ale căror ecuații sînt date în (2).

Din antichitate, matematicienii știau că multe curbe interesante pot fi definite și desenate cu ajutorul unor instrumente mecanice simple. Dintre aceste „curbe mecanice", *cicloidele* sînt dintre cele mai remarcabile. Ptolomeu (aproximativ 200 î.e.n) le folosea într-un mod foarte ingenios, pentru a descrie mișcările planetelor în spațiu.

Cea mai simplă cicloidă este curba descrisă de un punct fix de pe circumferința unui cerc, care se rostogolește fără alunecare de-a lungul unei drepte. Fig. 53 arată patru poziții ale punctului P de pe cercul care se rostogolește. Aspectul general al cicloidei este acela al unui șir de arcuri care se sprijină pe dreaptă.

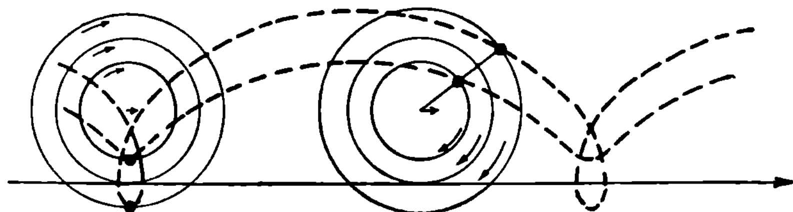


Fig. 54. Cicloide generale

Se pot obține modificări ale acestei curbe, alegînd punctul P, fie în interiorul unui cerc (ca și cum s-ar afla pe spița unei roți), sau pe prelungirea unei raze, dincolo de marginea discului. Fig. 54 ilustrează aceste două curbe.

Putem obține o altă modificare a cicloidei, permițînd cercului să se rostogolească nu în lungul unei drepte, ci pe un alt cerc. Dacă cercul c care se rostogolește, de rază r, rămîne tangent interior unui cerc C de rază $R > r$, locul geometric descris de un punct fix de pe circumferința lui c se numește *hipocicloidă*.

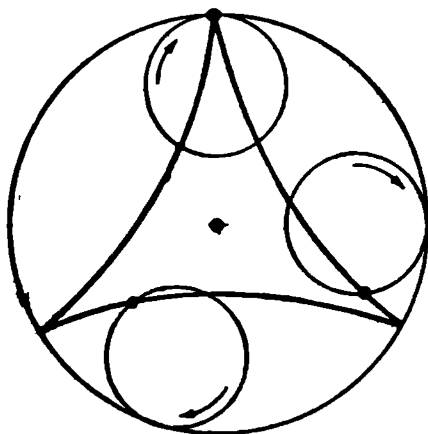


Fig. 55. Hipocicloida cu trei puncte de întoarcere

Dacă cercul c descrie întreaga circumferință a lui C o singură dată, atunci punctul P se va întoarce în poziția inițială numai dacă raza lui C este un multiplu întreg al razei lui c. Fig. 55 ilustrează cazul în care $R = 3r$. Mai general, dacă raza lui C este de m/n ori mai mare decît cea a lui c, hipocicloida se va închide după n circuite în lungul lui C, și va consta din m arcuri. Un caz particular interesant apare dacă $R = 2r$. Orice punct P al cercului interior va descrie atunci un diametru al cercului mai mare (fig. 56). Propunem cititorului ca problemă demonstrarea acestui lucru.

Un alt tip de cicloidă poate fi generat cu ajutorul unui cerc care se rostogolește, rămînînd tangent exterior unui cerc dat. O astfel de curbă se numește *epicicloidă*.

*4. Articulații. Inversorii lui Peaucellier și Hart

Renunțăm pentru moment la subiectul cicloidelor (ele vor apare din nou într-un loc neașteptat), pentru a considera alte metode de generare a curbelor. Cele mai simple instrumente mecanice pentru trasarea curbelor sînt *articulațiile*. O articulație constă dintr-un număr de bare rigide, legate într-un anumit

fel în puncte de legătură mobile, astfel încît întregul sistem să aibă suficientă libertate pentru a permite unui punct al său să descrie o anumită curbă. Compasul este în realitate o articulaţie simplă, formată, în principiu, dintr-o singură bară, fixată într-un punct.

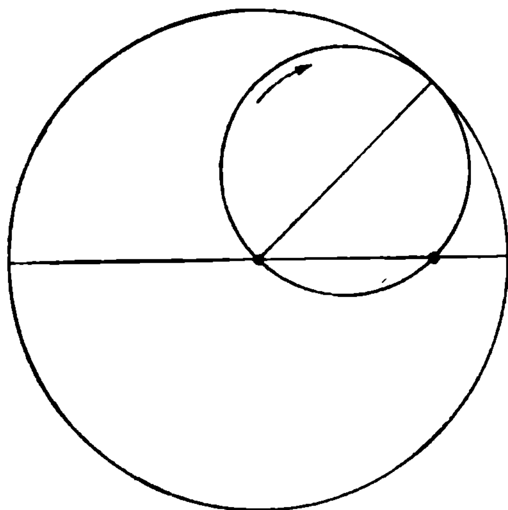


Fig. 56. Mişcarea rectilinie a punctelor care se află pe un cerc care se rostogoleşte în interiorul unui cerc de rază dublă

Articulaţiile au fost utilizate de multă vreme la construirea maşinilor. Unul dintre exemplele celebre din punct de vedere istoric, ,,paralelogramul lui Watt'', a fost inventat de James Watt, pentru rezolvarea problemei legării pistonului maşinii sale cu aburi de un punct de pe volan, în aşa fel încît rotaţia volanului să provoace mişcarea pistonului în lungul unei drepte. Soluţia lui Watt a fost doar aproximativă şi, în ciuda eforturilor făcute de mulţi matematicieni distinşi, problema construirii unei articulaţii pentru a deplasa un punct *cu exactitate* în lungul unei drepte a rămas nerezolvată. Pe vremea cînd demonstraţiile pentru imposibilitatea rezolvării anumitor probleme atrăgeau atenţia generală, s-a emis ipoteza că construirea unei astfel de articulaţii ar fi imposibilă. A fost o mare surpriză atunci cînd, în 1864, un ofiţer de marină francez, Peaucellier, a inventat o articulaţie simplă, care rezolva problema. Însă o dată cu introducerea unor lubrifianţi eficienţi, problema tehnică şi-a pierdut însemnătatea pentru maşinile cu aburi.

Scopul articulației lui Peaucellier este de a transforma mișcarea circulară în mișcare rectilinie. Ea se bazează pe teoria inversiunii, discutată în §4. După cum se arată în fig. 58, articulația constă din șapte bare rigide: două de lungime t, patru de lungime s, iar a șaptea de lungime arbitrară. O și R sînt

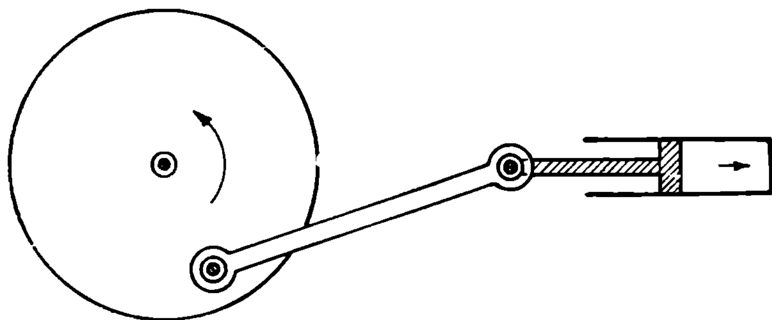


Fig. 57. Transformarea mișcării rectilinii în mișcare de rotație

două puncte fixe, așezate astfel încît $OR = PR$. Întregul aparat se poate mișca cu anumite restricții. Vom demonstra că *dacă* P *descrie un arc de cerc cu centrul în* R *și de rază* PR, *atunci* Q *descrie un segment de dreaptă.* Notînd cu T piciorul perpendicularei coborîtă din S pe OQ, observăm că

$$OP \cdot OQ = (OT - PT)(OT + PT) = OT^2 - PT^2 = (OT^2 + ST^2) -$$
$$- (PT^2 + ST^2) = t^2 - s^2.$$

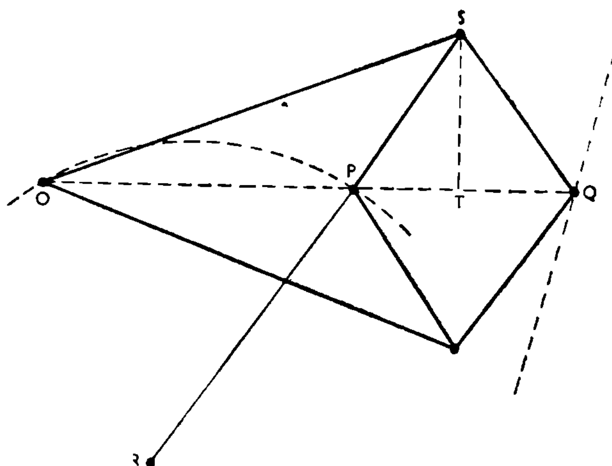


Fig. 58. Transformarea lui Peaucellier a mișcării de rotație în mișcare rectilinie

Cantitatea $t^2 - s^2$ este o constantă, pe care o notăm cu r^2. Deoarece $OP \cdot OQ = r^2$, P și Q sînt puncte inverse în raport cu un cerc de rază r și de centru O. Cînd P descrie o traiectorie circulară (care trece prin O), Q descrie curba inversă cercului. Această curbă trebuie să fie o dreaptă, pentru că am demonstrat că inversul unui cerc care trece prin O este o dreaptă. Astfel, traiectoria lui Q este o dreaptă, trasată fără a utiliza rigla.

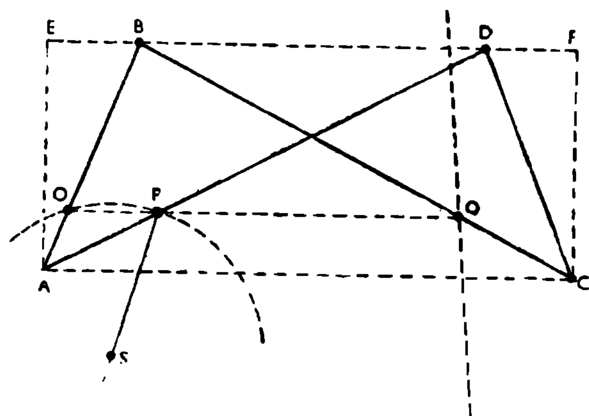


Fig. 59. Inversorul lui Hart

O altă articulație care rezolvă aceeași problemă este inversorul lui Hart. Acesta constă din cinci bare legate, după cum se arată în fig. 59. Aici $AB = CD$, $BC = AD$, iar O, P și Q sînt puncte fixe, care se află respectiv pe barele AB, AD, CB, astfel încît $AO/OB = AP/PD = CQ/QB = m/n$. Punctele O și S sînt fixate în plan, astfel încît $OS = PS$, în timp ce restul articulației se poate mișca în voie. Evident, AC este întotdeauna paralelă cu BD. Deci O, P și Q sînt coliniare și OP este paralelă cu AC. Coborîți AE și CF perpendiculare pe BD. Avem

$$AC \cdot BD = EF \cdot BD = (ED + EB)(ED - EB) = ED^2 - EB^2.$$

Însă

$$ED^2 + AE^2 = AD^2, \quad \text{și} \quad EB^2 + AE^2 = AB^2.$$

Deci

$$ED^2 - EB^2 = AD^2 - AB^2.$$

Însă

$$OP/BD = AO/AB = m/(m + n) \qquad \text{și} \qquad OQ/AC = OB/AB = n/(m + n).$$

Astfel

$$OP \cdot OQ = [mn/(m+n)^2]BD \cdot AC = [mn/(m+n)^2](AD^2 - AB^2).$$

Această cantitate rămîne neschimbată, oricare ar fi poziția posibilă a articulației. De aceea, P și Q sînt puncte inverse în raport cu un cerc cu centrul în O. Cînd mișcăm articulația, P descrie un cerc cu centrul în S, care trece prin O, în timp ce inversul său Q descrie o dreaptă.

Se pot construi și alte articulații (cel puțin în principiu) care trasează elipse, hiperbole și orice curbă dată printr-o ecuație algebrică $f(x,y) = 0$, de orice grad.

§ 6. DIN NOU DESPRE INVERSIUNE ȘI APLICAȚIILE EI

1. Invarianța unghiurilor. Familii de cercuri

Cu toate că inversiunea față de un cerc modifică foarte mult aspectul figurilor geometrice, este remarcabil faptul că noile figuri posedă multe din proprietățile figurii inițiale. Acestea sînt proprietățile care rămîn neschimbate, adică sînt „invariante" prin transformare. După cum știm, inversiunea transformă cercurile și dreptele în cercuri și drepte. Adăugăm acum o altă proprietate importantă: *Unghiul dintre două drepte sau două curbe este invariant prin transformare.* Prin aceasta înțelegem că oricare două curbe care se intersectează se transformă prin inversiune în alte două curbe, care se intersectează sub același unghi. Prin unghi dintre două curbe înțelegem, desigur, unghiul dintre tangentele lor.

Demonstrația poate fi înțeleasă din fig. 60, care ilustrează cazul particular al unei curbe C, care intersectează o dreaptă OL într-un punct P. Inversul C' al lui C întîlnește pe OL în punctul invers P' care, dat fiind că OL este propria sa inversă, se află pe OL. Vom arăta că unghiul x_0 dintre OL și tangenta la C în P este egal în mărime cu unghiul corespunzător y_0. Pentru a face aceasta alegem un punct A pe curba C, lîngă P, și ducem secanta AP. Inversul lui A este un punct A', care, fiind atît pe dreapta, OA, cît și pe curba C', trebuie să fie la intersecția lor. Ducem secanta $A'P'$. Din definiția inversiunii, avem

$$r^2 = OP \cdot OP' = OA \cdot OA'$$

sau

$$\frac{OP}{OA} = \frac{OA'}{OP'},$$

adică triunghiurile OAP și $OA'P'$ sînt asemenea. Prin urmare, unghiul x este egal cu unghiul $OA'P'$, pe care îl notăm cu y. Pasul final constă în a face pe A

să se deplaseze în lungul lui C, apropiindu-l de punctul P. Acest lucru face ca secanta AP să tindă spre tangenta în P la curba C, în timp ce unghiul x tinde spre x_0. În același timp, A' se va apropia de P', $A'P'$ va tinde spre tangenta în P' și unghiul y tinde spre y_0. Deoarece $x = y$ în orice poziție a lui A, trebuie să avem la limită $x_0 = y_0$.

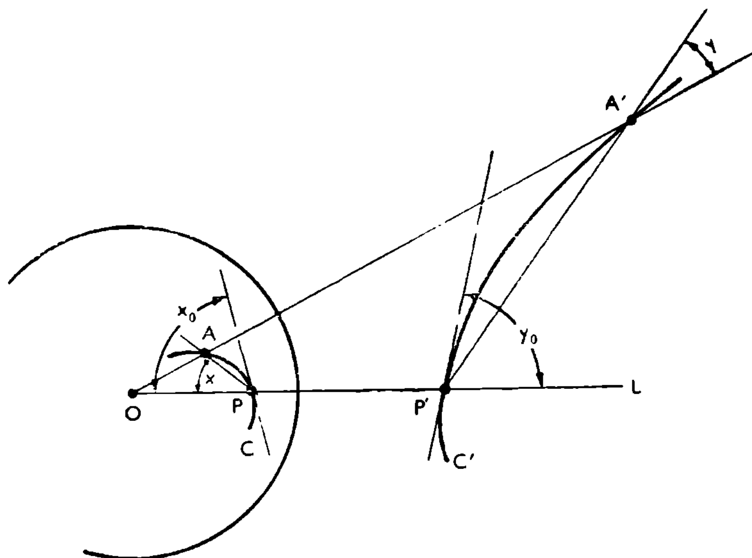


Fig. 60. Invarianța unghiurilor prin inversiune

Demonstrația dată este doar în parte completă, pentru că am considerat numai cazul unei curbe care intersectează o dreaptă ce trece prin O. Cazul general a două curbe C, C', care formează un unghi z în P, se tratează acum cu ușurință. Într-adevăr, este evident faptul că dreapta OPP' împarte unghiul z în două părți, fiecare fiind conservată prin inversiune.

Trebuie să remarcăm că deși inversiunea conservă *mărimea* unghiurilor, ea schimbă *sensul* lor; adică dacă o rază dusă prin P parcurge unghiul x_0 în sensul opus acelor unui ceasornic, imaginea ei va parcurge unghiul y_0 în sensul acelor unui ceasornic.

O consecință particulară a invariației unghiului prin inversiune este faptul că două cercuri sau drepte ortogonale (care se intersectează sub unghiuri drepte) rămîn ortogonale după inversiune, în timp ce două cercuri tangente, care se intersectează după un unghi nul, rămîn tangente.

Să considerăm familia tuturor cercurilor care trec prin centrul de inversiune O și printr-un alt punct fix A din plan. Din § 4, secțiunea 2, știm că această

familie de cercuri se transformă într-o familie de drepte care trec prin A', imaginea lui A. Familia de cercuri ortogonale familiei inițiale se transformă în cercuri ortogonale dreptelor care trec prin A', după cum se arată în fig. 61. (Cercurile ortogonale sînt trasate cu linie întreruptă.) Figura simplă a dreptelor care

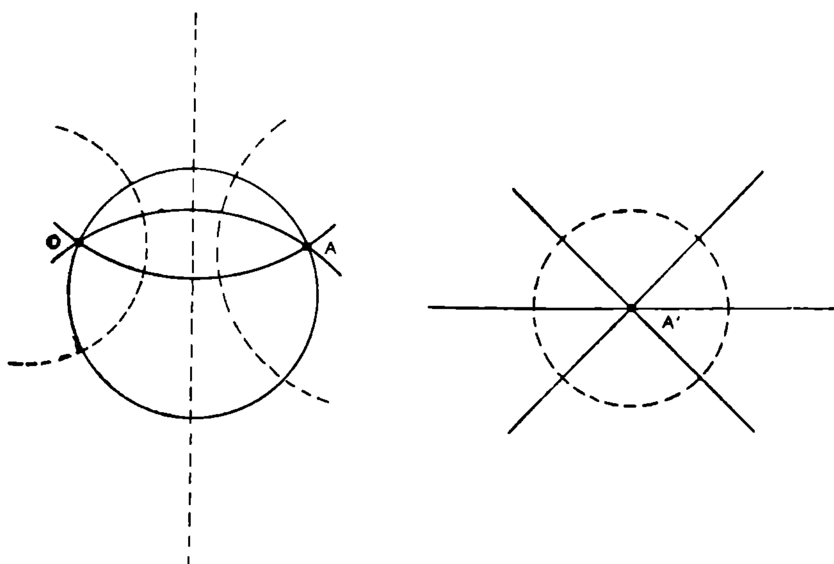


Fig. 61. Două sisteme de cercuri ortogonale legate prin inversiune

pornesc dintr-un punct pare a fi cu totul deosebită de aceea a cercurilor și, cu toate acestea, vedem că ele sînt strîns legate, iar din punctul de vedere al teoriei inversiunii, ele sînt întru totul echivalente.

Un alt exemplu al efectului inversiunii este dat de o familie de cercuri tangente în polul de inversiune. Prin transformare ele devin un sistem de drepte paralele pentru că imaginile cercurilor sînt drepte care nu se intersectează două cîte două, deoarece cercurile inițiale se întîlnesc doar în O.

2. Aplicație la problema lui Apollonius

O bună ilustrare a utilității teoriei inversiunii este următoarea rezolvare geometrică simplă a problemei lui Apollonius. Prin inversiune în raport cu un pol oarecare, problema lui Apollonius pentru trei cercuri date poate fi transformată în problema corespunzătoare pentru alte trei cercuri. Prin urmare, dacă putem rezolva problema pentru un sistem format din trei cercuri, atunci ea este rezolvată pentru orice alt triplet de cercuri, obținut din primul prin

inversiune. Vom utiliza acest fapt, alegînd dintre toate tripletele echivalente de cercuri unul pentru care problema este aproape banală.

Vom începe cu trei cercuri cu centrele în A, B, C şi vom presupune că cercul cerut U, cu centrul O şi raza ρ este tangent exterior celor trei cercuri date. Dacă mărim razele celor trei cercuri cu aceeaşi cantitate d, atunci cercul cu

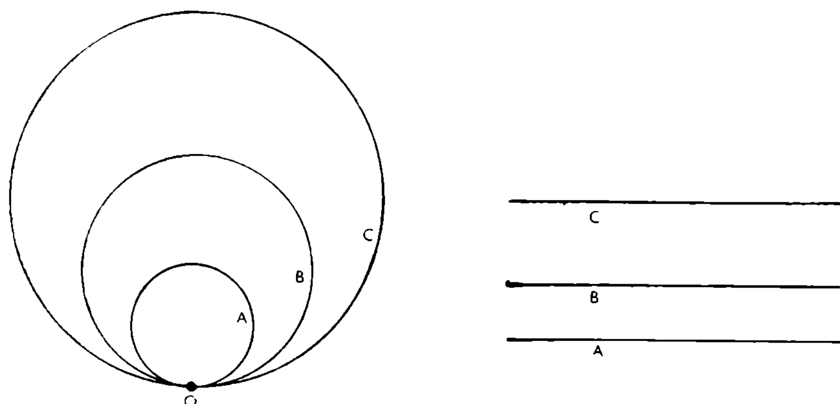


Fig. 62. Transformarea cercurilor tangente în drepte paralele

acelaşi centru O şi raza $\rho - d$ va rezolva noua problemă. Ca pregătire, vom folosi acest lucru, pentru a înlocui cele trei cercuri date cu alte trei, astfel încît două dintre ele să fie tangente într-un punct K (fig. 63). Apoi inversăm întreaga

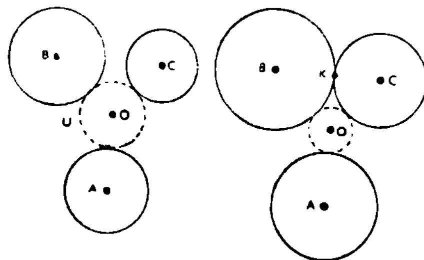


Fig. 63. Construcţie preliminară rezolvării problemei lui Apollonius

figură în raport cu un cerc cu centrul în K. Cercurile cu centrele în B şi C devin drepte paralele b şi c, în timp ce al treilea cerc se transformă într-un cerc a (fig. 64). Ştim că a, b, c pot fi construite toate cu ajutorul riglei şi compasu-

lui. Cercul necunoscut se transformă într-un cerc u, tangent cercurilor a, b, c. Raza sa r este egală, evident, cu jumătatea distanței dintre b și c. Centrul său O' este una dintre cele două intersecții ale dreptei care trece la mijlocul distanței dintre b și c, cu cercul de centru A' (centrul lui a) și de rază $r + s$ (s

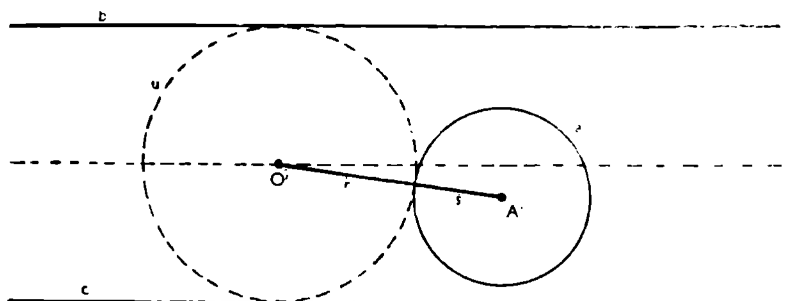


Fig. 64. Rezolvarea problemei lui Apollonius

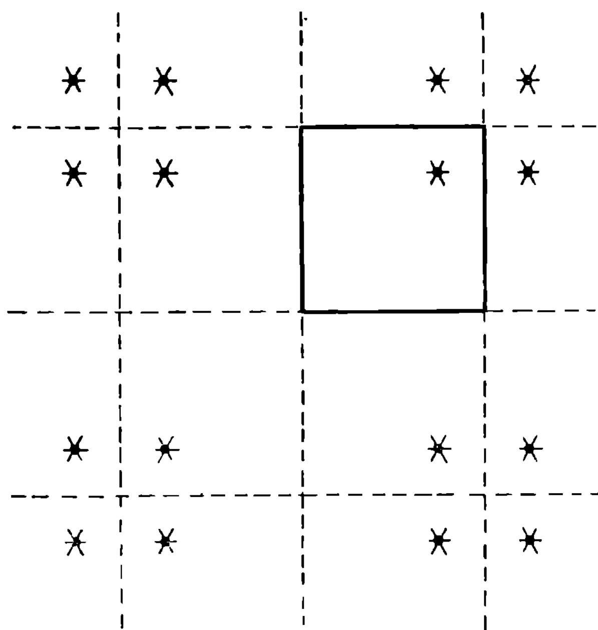


Fig. 65. Reflexii iterate în pereți dreptunghiulari

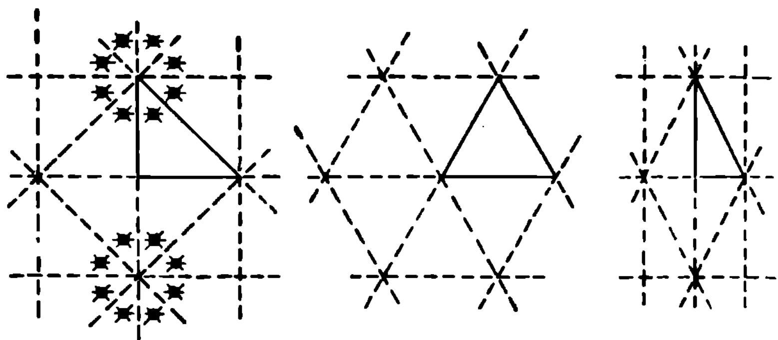


Fig. 66. Constelații regulate de oglinzi triunghiulare

Fig. 67. Reflexii iterate în sisteme formate din două cercuri

fiind raza lui *a*). În sfîrşit, construind cercul invers lui *u* găsim centrul cercului lui Apollonius *U*. (Centrul său *O* va fi inversul în raport cu cercul de inversiune al punctului invers lui *K* în raport cu *u*.)

Oricine cunoaşte fenomenele de reflexie iterată, care apar atunci cînd se utilizează mai multe oglinzi. Dacă cei patru pereţi ai unei camere dreptunghiulare ar fi acoperiţi cu oglinzi ideale, neabsorbante, un punct luminos ar avea o

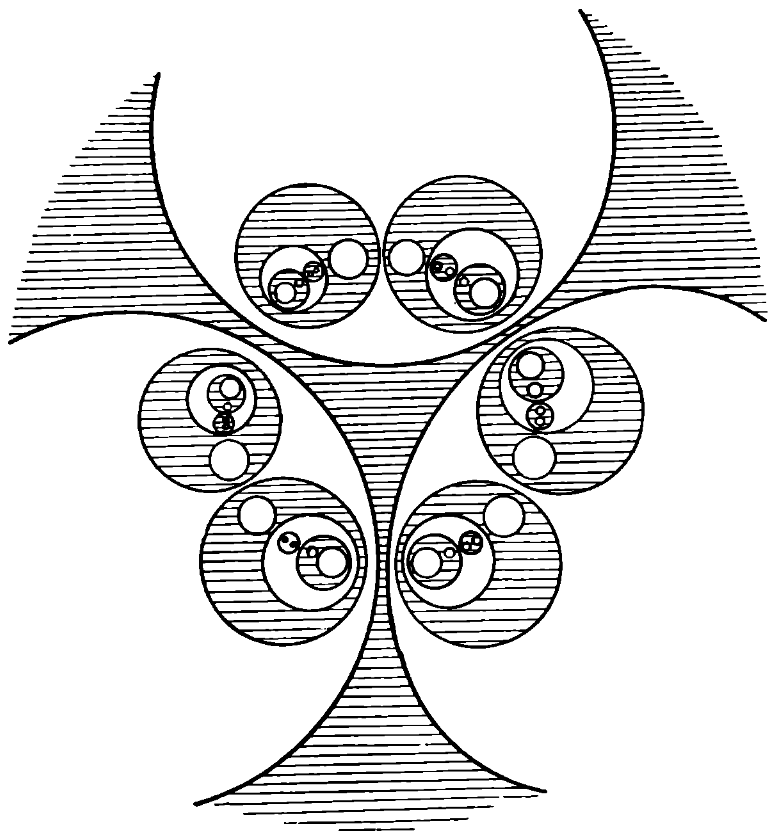


Fig. 68. Reflexia într-un sistem format din trei cercuri

infinitate de imagini, fiecărei camere egale obţinută prin reflexie corespunzîndu-i cîte unul (fig. 65). O constelaţie mai puţin regulată de oglinzi, de exemplu, trei oglinzi, dă o serie mult mai complicată de imagini. Configuraţia care rezultă poate fi descrisă cu uşurinţă numai dacă triunghiurile reflectate formează o acoperire fără suprapuneri a planului. Aceasta se întîmplă numai

în cazul triunghiului dreptunghic isoscel, al triunghiului echilateral și al jumătății dreptunghice a acestuia din urmă (fig. 66).

Situația devine mult mai interesantă dacă considerăm inversiuni iterate într-o pereche de cercuri. Stînd între două oglinzi circulare concentrice, am vedea o infinitate de alte cercuri concentrice cu ele. Un șir din aceste cercuri tinde spre infinit, în timp ce altul tinde spre centru. Cazul a două cercuri exterioare este puțin mai complicat. Aici cercurile și imaginile lor se reflectă succesiv unul în altul, devenind din ce în ce mai mici cu fiecare reflexie, pînă ce ele se reduc la două puncte, cîte unul în fiecare cerc. (Aceste puncte au proprietatea de a fi reciproc inverse în raport cu cele două cercuri.) Toate acestea sînt arătate în fig. 67. Folosirea a trei cercuri ne conduce la modelul frumos indicat în fig. 68.

GEOMETRIA PROIECTIVĂ. AXIOMATICA. GEOMETRII NEEUCLIDIENE

§ 1. INTRODUCERE

1. Clasificarea proprietăților geometrice. Invarianța prin transformări

Geometria se ocupă cu studiul proprietăților figurilor din plan sau spațiu. Aceste proprietăți sînt atît de numeroase și de variate, încît este necesar un principiu de clasificare pentru a face ordine în această bogăție de cunoștințe. De exemplu, am putea introduce o clasificare bazată pe metoda folosită la deducerea teoremelor. Din acest punct de vedere, se face de obicei o distincție între procedeele „sintetic" și „analitic". Primul dintre acestea este metoda axiomatică clasică a lui Euclid, în care subiectul este tratat pe baze pur geometrice, independent de algebră și de conceptul de continuu numeric, și în care teoremele sînt deduse prin raționament logic dintr-un corp inițial de propoziții, numite axiome sau postulate. A doua metodă se bazează pe introducerea coordonatelor numerice și folosește tehnica algebrei. Această metodă a adus o modificare profundă în matematică, care a avut ca rezultat o unificare a geometriei, analizei și algebrei într-un sistem organic.

În acest capitol, clasificarea bazată pe metodă va fi mai puțin importantă decît o clasificare bazată pe *conținut*, pe caracterul teoremelor, independent de metodele folosite pentru demonstrarea lor. În geometria plană elementară distingem între teoremele care se referă la congruența figurilor, care utilizează noțiunile de lungime și unghi, și teoremele care se referă la asemănarea figurilor, care folosesc doar noțiunea de unghi. Această distincție particulară nu este foarte importantă, deoarece lungimea și unghiul sînt atît de strîns legate, încît este oarecum artificial să le separăm. (Studiul acestei legături constituie obiectul principal al trigonometriei.) În schimb, putem spune că în geometria elementară avem de-a face cu *mărimi* — lungimi, unghiuri și arii. Două figuri sînt echivalente din acest punct de vedere, dacă ele sînt *congruente*, adică dacă una poate fi obținută din cealaltă printr-o *mișcare rigidă*, care schimbă numai poziția figurii, dar nu și mărimea ei. Acum se pune problema dacă noțiunea de mărime și noțiunile derivate de congruență și asemă-

nare sînt esenţiale în geometrie sau dacă nu cumva figurile geometrice au proprietăţi mai adînci, care se păstrează chiar prin transformări mai generale decît mişcările rigide. Vom vedea că astfel de proprietăţi există într-adevăr.

Să presupunem că trasăm un cerc şi o pereche de diametri perpendiculari pe o bucată dreptunghiulară de lemn moale, ca în fig. 69. Dacă aşezăm această bucată între fălcile unei menghine puternice şi dacă o comprimăm, reducînd-o

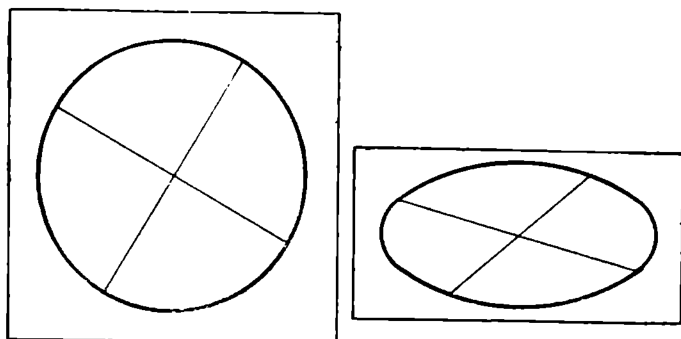


Fig. 69. Comprimarea unui cerc

la jumătatea lăţimii iniţiale, cercul va deveni o elipsă, iar unghiurile dintre diametrii elipsei nu vor mai fi drepte. Cercul are proprietatea că toate punctele sale sînt echidistante faţă de centru, în timp ce elipsa nu are o astfel de proprietate. Astfel, ar putea să pară că toate proprietăţile geometrice ale configuraţiei originale sînt distruse prin comprimare. Dar aceasta nu este adevărat ; de exemplu, propoziţia că centrul înjumătăţeşte fiecare diametru este adevărată atît la cerc, cît şi la elipsă : în cazul de faţă avem o astfel de proprietate a figurii care se păstrează chiar după o modificare drastică a mărimilor figurii originale. Această observaţie sugerează posibilitatea clasificării teoremelor referitoare la o figură geometrică, după cum ele rămîn, sau nu, adevărate cînd figura este supusă unei comprimări uniforme. Mai general, dată fiind o anumită clasă de transformări ale unei figuri (ca de pildă clasa tuturor mişcărilor rigide, a comprimărilor, a inversiunilor în raport cu cercuri etc.), ne putem întreba care sînt proprietăţile figurii care rămîn neschimbate prin transformările clasei considerate. Totalitatea teoremelor referitoare la aceste proprietăţi va forma *geometria asociată acestei clase de transformări*. Ideea clasificării diferitelor ramuri ale geometriei, potrivit claselor de transformări considerate a fost propusă de Felix Klein (1849—1925) în celebra sa cuvîntare *Programul de la*

Erlangen, ținută în 1872. De atunci, această idee a influențat în mare măsură gîndirea geometrică.

În cap. V vom descoperi faptul foarte surprinzător că anumite proprietăți ale figurilor geometrice sînt atît de profunde, încît ele persistă chiar după ce figurile sînt supuse unor deformări arbitrare; astfel figurile desenate pe o bucată de cauciuc, care este deformat într-un mod arbitrar, își mențin cîteva din caracteristicile lor inițiale. În acest capitol ne vom ocupa însă de acele proprietăți care rămîn neschimbate sau „invariante" printr-o anumită clasă de transformări, care se află între clasa foarte restrînsă a mișcărilor rigide, pe de o parte, și clasa cea mai generală a deformărilor arbitrare, pe de alta. Aceasta este clasa „transformărilor proiective".

2. Transformări proiective

Studiul acestor proprietăți geometrice a fost impus matematicienilor cu mult în urmă, de problemele de *perspectivă*, care au fost studiate de artiști, ca de pildă Leonardo da Vinci și Albrecht Dürer. Imaginea desenată de un pictor trebuie privită ca proiecție a originalului pe suprafața pînzei, centrul de proiecție fiind ochiul pictorului. Prin proiecție, lungimile și unghiurile sînt în mod necesar deformate într-un mod care depinde de pozițiile relative ale diferitelor obiecte pictate. Totuși, de obicei, structura geometrică a originalului poate fi recunoscută pe pînză. Cum este posibil acest lucru? El se datorește faptului că există proprietăți geometrice „invariante prin proiecție", proprietăți care rămîn neschimbate în imagine și fac posibilă identificarea. Găsirea și analiza acestor proprietăți reprezintă obiectul geometriei proiective.

Este limpede că în această ramură a geometriei nu pot fi propoziții referitoare la lungimi și unghiuri sau despre congruențe. Unele fapte izolate referitoare la proprietățile proiective erau cunoscute încă din secolul al XVII-lea și uneori, ca în cazul „teoremei lui Menelaus", chiar din antichitate. Dar un studiu sistematic al geometriei proiective a fost făcut pentru prima dată la sfîrșitul secolului al XVIII-lea, cînd École Polytechnique din Paris a inițiat o nouă pagină în istoria matematicii, în special în geometrie. Această școală, produs al revoluției franceze, a dat mulți ofițeri pentru serviciile militare ale republicii. Unul dintre absolvenții ei a fost J. V. Poncelet (1788—1867), care a scris celebrul său *Traité des propriétés projectives des figures* în 1813, pe cînd era prizonier de război în Rusia. În secolul al XIX-lea, sub influența lui Steiner, von Staudt, Chasles și a altora, geometria proiectivă a devenit unul dintre principalele subiecte ale cercetării matematice. Popularitatea ei se datora, pe de o parte, marelui ei farmec estetic, iar pe de altă parte, efectului ei clarificator asupra geometriei în ansamblu și legăturii ei intime cu geometria neeuclidiană și algebra.

§ 2. NOŢIUNI FUNDAMENTALE

1. Grupul transformărilor proiective

Definim mai întîi clasa sau „grupul"[1] transformărilor proiective. Să presupunem că avem în spaţiu două plane π şi π', nu neapărat paralele. Atunci putem efectua o *proiecţie centrală* a lui π pe π' dintr-un centru O, care nu se află în π sau în π', definind imaginea fiecărui punct P din π, ca fiind punctul P' din

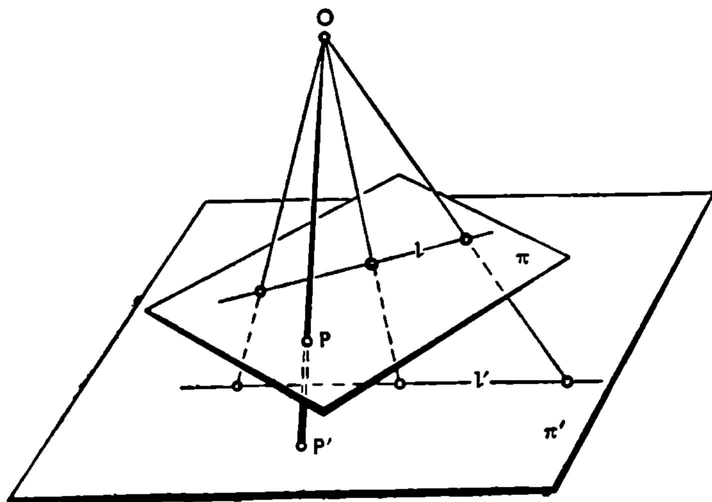


Fig. 70. Proiecţia centrală

π', astfel încît P şi P' să se afle pe aceeaşi dreaptă dusă din O. De asemenea, putem efectua o *proiecţie paralelă*, în care liniile de proiecţie sînt toate paralele. În acelaşi mod putem defini proiecţia unei drepte l dintr-un plan π pe o altă dreaptă l' din π', dintr-un punct O, sau printr-o proiecţie paralelă.

Orice aplicaţie a unei figuri pe alta, printr-o proiecţie centrală sau paralelă sau printr-o succesiune finită de astfel de proiecţii, se numeşte *transformare proiectivă*[2]. *Geometria proiectivă* a planului sau a dreptei constă din ansamblul

[1] Termenul „grup" aplicat unei clase de transformări implică faptul că aplicarea succesivă a două transformări ale clasei este tot o transformare a aceleiaşi clase şi că „inversul" unei transformări a clasei aparţine din nou clasei. Proprietăţile grupale ale operaţiilor matematice au jucat şi joacă un rol foarte mare în multe domenii, cu toate că în geometrie, poate, importanţa noţiunii de grup a fost puţin exagerată.

[2] Două figuri legate printr-o *singură* proiecţie se numesc *perspective*. Astfel, o figură F este legată de figura F' printr-o transformare proiectivă, dacă F şi F' sînt perspective sau dacă putem găsi o succesiune de figuri, $F, F_1, F_2, \ldots, F_n, F'$, astfel încît fiecare figură să fie perspectivă cu următoarea.

propozițiilor geometrice, care nu sînt modificate prin transformări proiective arbitrare ale figurilor la care se referă. Spre deosebire de aceasta, vom numi *geometrie metrică* ansamblul acelor propoziții, care se referă la mărimile figurilor, invariante numai în raport cu clasa mișcărilor rigide.

Unele proprietăți proiective pot fi recunoscute imediat. Un punct, desigur, se proiectează într-un punct. Mai mult *o dreaptă se proiectează într-o dreaptă*;

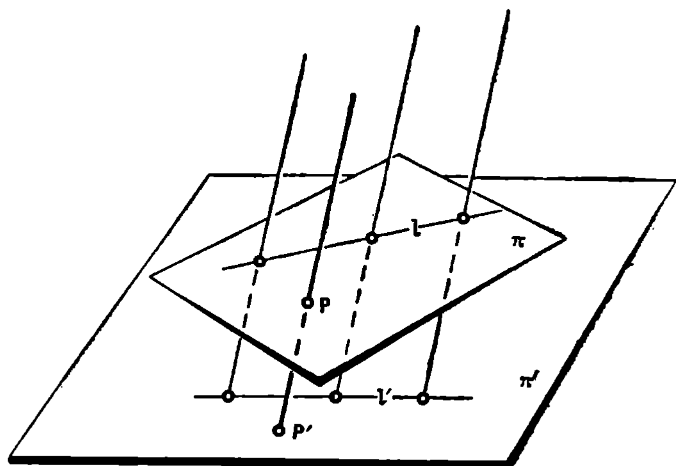


Fig. 71. Proiecția paralelă

pentru că dacă dreapta l din π este proiectată pe planul π', intersecția lui π' cu planul care trece prin O și l va fi dreapta l'[3]. Dacă un punct A și o dreaptă l sînt incidente[4], atunci după orice proiecție, punctul corespunzător A' și dreapta l' vor fi și ele incidente. Astfel, *incidența unui punct cu o dreaptă este invariantă prin grupul proiectiv*. Din acest fapt rezultă multe consecințe simple, dar importante. Dacă trei sau mai multe puncte sînt *coliniare*, adică sînt incidente cu o dreaptă, atunci și imaginile lor vor fi coliniare. De asemenea, dacă în planul π trei sau mai multe drepte sînt *concurente*, adică incidente cu un punct, atunci imaginile lor vor fi de asemenea drepte concurente. În timp ce aceste proprietăți simple — incidența, coliniaritatea și concurența — sînt *proprietăți proiective* (adică proprietăți invariante prin proiecții),

[3] Există excepții în cazul în care dreapta OP, sau planul care trece prin O și prin l este paralel cu planul π'. Vom îndepărta aceste excepții în § 4.

[4] Un punct și o dreaptă se numesc *incidente* dacă dreapta trece prin punct sau dacă punctul se află pe dreaptă. Cuvîntul neutru folosit subliniază reciprocitatea relației considerate.

măsurile lungimilor şi unghiurilor, ca şi rapoartele acestor mărimi, de obicei sînt modificate prin proiecţie. Triunghiurile isoscele sau echilaterale se pot proiecta în triunghiuri ale căror laturi sînt toate diferite. Deci cu toate că „triunghiul" este o noţiune a geometriei proiective, „triunghiul echilateral" nu mai este o astfel de noţiune şi aparţine doar geometriei metrice.

2. Teorema lui Desargues

Una dintre primele descoperiri în geometria proiectivă este celebra teoremă a lui Desargues (1593—1662). *Dacă într-un plan, două triunghiuri* ABC *şi* A'B'C' *sînt astfel situate încît dreptele care unesc vîrfurile corespunzătoare să fie concurente într-un punct* O, *atunci laturile corespunzătoare, prelungite, se intersectează în trei puncte coliniare.* Figura 72 ilustrează teorema, iar cititorul ar trebui să deseneze alte figuri pentru a o verifica experimental. Demonstraţia teoremei nu este banală, în ciuda simplităţii figurii în care intervin numai drepte. Teorema aparţine evident geometriei proiective, deoarece dacă proiectăm întreaga figură pe un alt plan, ea îşi va păstra proprietăţile din enunţul teoremei. Vom reveni la această teoremă la p.206. Pentru moment, dorim să atragem atenţia asupra faptului remarcabil că teorema lui Desargues este adevărată şi dacă cele două triunghiuri se află în două plane *diferite* (neparalele) şi că această teoremă a lui Desargues din geometria tridimensională se demon-

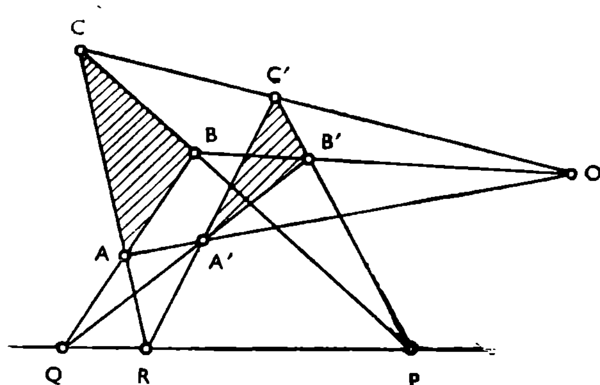


Fig. 72. Configuraţia lui Desargues în plan

strează foarte uşor. Să presupunem că dreptele AA', BB' şi CC' se intersectează în O (fig. 73), potrivit ipotezei. Atunci AB se află în acelaşi plan cu $A'B'$, astfel încît aceste două drepte se intersectează într-un punct Q; de asemenea, AC şi $A'C'$ se intersectează în R, iar BC şi $B'C'$ se intersectează în P. Deoarece P, Q şi R se află pe prelungirile laturilor triunghiurilor ABC şi $A'B'C'$, ele se

află în acelaşi plan cu fiecare dintre aceste două triunghiuri şi deci trebuie să se afle pe dreapta de intersecţie a acestor două plane. De aceea, P, Q şi R sînt coliniare, după cum trebuia demonstrat.

Această demonstraţie simplă sugerează faptul că am putea demonstra teorema bidimensională, printr-o aşa-numită trecere la limită, făcînd ca întreaga

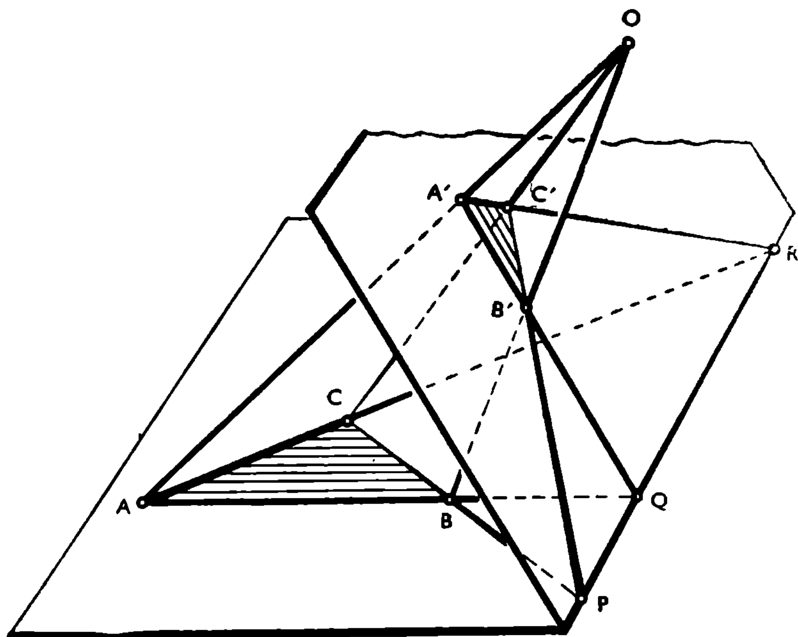


Fig. 73. Configuraţia lui Desargues în spaţiu

figură să se turtească pînă ce cele două plane coincid şi punctul O, împreună cu toate celelalte, devin coplanare. Există însă o oarecare dificultate la efectuarea acestei treceri la limită, pentru că dreapta de intersecţie PQR nu este determinată în mod unic, atunci cînd cele două plane coincid. Însă configuraţia reprezentată în fig. 72 poate fi privită ca schiţă perspectivă a configuraţiei spaţiale a fig. 73 şi acest fapt poate fi folosit pentru a demonstra teorema în cazul planului.

În realitate există o deosebire fundamentală între teorema lui Desargues din plan şi cea din spaţiu. Demonstraţia tridimensională pe care am prezentat-o folosea un raţionament geometric bazat numai pe noţiunile de incidenţă şi intersecţie a punctelor, dreptelor şi planelor. Se poate arăta că demonstraţia teoremei bidimensionale, *dacă trebuie efectuată fără a ieşi din*

plan, necesită folosirea noțiunii de asemănare a figurilor, care este bazată pe noțiunea metrică de lungime și nu mai este o noțiune proiectivă.

Reciproca teoremei lui Desargues afirmă că dacă ABC și $A'B'C'$ sînt două triunghiuri situate astfel, încît punctele de intersecție ale laturilor corespunzătoare să fie coliniare, atunci dreptele care unesc vîrfurile corespunzătoare sînt concurente. Lăsăm pe seama cititorului, ca exercițiu, demonstrarea ei, în cazul în care cele două triunghiuri se află în plane neparalele.

§ 3. BIRAPORTUL

1. Definiția și demonstrarea invarianței

Tot așa cum lungimea unui segment de dreaptă este cheia geometriei metrice, există o noțiune fundamentală a geometriei proiective, în raport cu care se pot exprima toate proprietățile proiective ale figurilor.

Dacă trei puncte A, B, C se află pe o dreaptă, o proiecție va schimba, în general, nu numai distanțele AB și BC, dar și raportul AB/BC. De fapt, *orice*, trei puncte A, B, C de pe o dreaptă l pot fi transformate întotdeauna în oricare trei puncte A', B', C' de pe o altă dreaptă l', prin două proiecții succesive. Pentru a face acest lucru, putem roti dreapta l' în jurul punctului C', pînă ce ea ocupă o poziție l'', paralelă cu l (fig. 74). Apoi proiectăm dreapta l pe l'' printr-o proiecție paralelă cu dreapta care unește pe C cu C', definind trei puncte A'', B'' și C'' $(=C')$. Dreptele care unesc pe A', A'' și B', B'' se vor intersecta într-un punct O, pe care îl alegem ca centru al celei de-a doua proiecții. Aceste două proiecții realizează construcția dorită[5].

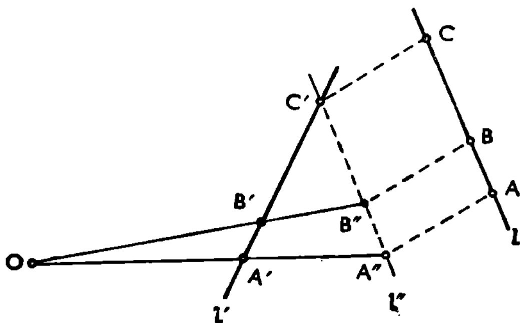


Fig. 74.

După cum am văzut, nici o cantitate care depinde numai de trei puncte de pe o dreaptă nu poate fi invariantă prin proiecție. Însă — și aceasta este descoperirea decisivă a geometriei proiective — dacă pe o dreaptă sînt date *patru*

[5] Ce se întîmplă dacă dreptele care unesc punctele A', A'' și B', B'' sînt paralele?

puncte A, B, C, D care prin proiecție trec în punctele A', B', C', D' de pe o altă dreaptă, atunci există o anumită cantitate, numită *biraportul* celor patru puncte, care își păstrează valoarea prin proiecție. Aici avem o proprietate matematică a unei mulțimi de patru puncte de pe o dreaptă, care nu se modifică prin proiecție și care poate fi regăsită în orice imagine a dreptei. Biraportul nu este nici lungimea, nici raportul a două lungimi, ci *raportul a două astfel de rapoarte*: dacă considerăm rapoartele CA/CB și DA/DB, atunci raportul lor

A B C D

$$x = \frac{CA}{CB} \bigg/ \frac{DA}{DB},$$

este, prin definiție, biraportul celor patru puncte A, B, C, D, luate în această ordine.

Să arătăm acum că *biraportul a patru puncte este invariant prin proiecție*, adică dacă A, B, C, D și A', B', C', D' sînt punctele corespunzătoare pe două drepte legate prin proiecție, atunci

$$\frac{CA}{CB} \bigg/ \frac{DA}{DB} = \frac{C'A'}{C'B'} \bigg/ \frac{D'A'}{D'B'}.$$

Demonstrația se face prin metode elementare. Reamintim că aria unui triunghi este egală cu 1/2 (baza × înălțimea) și este dată și de jumătatea produsului a două laturi, înmulțită cu sinusul unghiului cuprins între ele. Atunci avem din fig. 75

$$\text{aria } OCA = \frac{1}{2} h \cdot CA = \frac{1}{2} OA \quad OC \sin \widehat{COA}$$

$$\text{aria } ODA = \frac{1}{2} h \cdot DA = \frac{1}{2} OA \cdot OD \cdot \sin \widehat{DOA}$$

$$\text{aria } OCB = \frac{1}{2} h \cdot CB = \frac{1}{2} OB \cdot OC \sin \widehat{COB}$$

$$\text{aria } ODB = \frac{1}{2} h \cdot DB = \frac{1}{2} OB \cdot OD \sin \widehat{DOB}.$$

Rezultă că

$$\frac{CA}{CB}\Big/\frac{DA}{DB}=\frac{CA}{CB}\cdot\frac{DB}{DA}=\frac{OA\cdot OC\sin\widehat{COA}}{OB\cdot OC\sin\widehat{COB}}\cdot\frac{OB\cdot OD\sin\widehat{DOB}}{OA\cdot OD\sin\widehat{DOA}}=$$

$$=\frac{\sin\widehat{COA}}{\sin\widehat{COB}}\cdot\frac{\sin\widehat{DOB}}{\sin\widehat{DOA}}.$$

Deci biraportul punctelor A, B, C, D depinde numai de unghiurile subîntinse în O de segmentele care unesc punctele A, B, C, D. Deoarece aceste unghiuri sînt aceleaşi pentru oricare patru puncte A', B', C', D' în care punctele A, B, C, D pot fi proiectate din O, rezultă că biraportul rămîne neschimbat prin proiecţie.

Faptul că biraportul a patru puncte rămîne neschimbat printr-o proiecţie *paralelă* rezultă din proprietăţi elementare ale triunghiurilor asemenea. Demonstraţia este lăsată pe seama cititorului, ca exerciţiu.

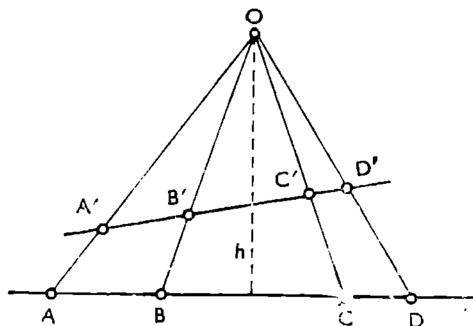


Fig. 75. Invarianţa biraportului prin proiecţie centrală

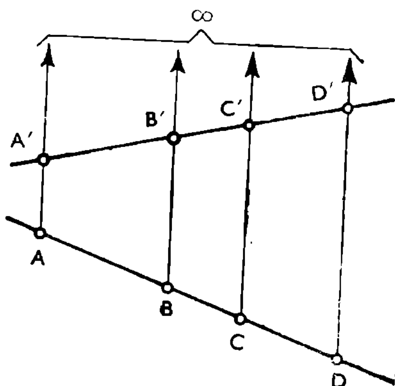


Fig. 76. Invarianţa biraportului prin proiecţie paralelă

Pînă acum am înţeles prin biraport a patru puncte A, B, C, D de pe o dreaptă l un raport în care intervin lungimi pozitive. Este mai comod să modificăm această definiţie în modul următor: alegem un sens pe l ca fiind cel pozitiv şi convenim ca lungimile măsurate în acest sens să fie pozitive, în timp ce lungimile

măsurate în sensul opus să fie negative. Atunci definim biraportul punctelor A, B, C D în această ordine, ca fiind cantitatea

$$(1) \qquad (A\ B\ C\ D) = \frac{AC}{CB} \Big/ \frac{DA}{DB},$$

în care numerele CA, CB, DA, DB sînt luate cu semnul lor. Deoarece o inversare a sensului pozitiv ales pe l schimbă doar semnul fiecărui termen al acestui raport, valoarea lui $(ABCD)$ nu va depinde de sensul ales. Se vede cu

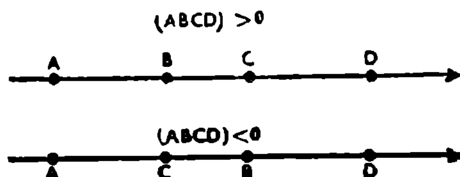


Fig. 77. Semnul biraportului

ușurință că $(ABCD)$ va fi negativ sau pozitiv, după cum perechea de puncte A, B este sau nu este separată de perechea CD. Deoarece această proprietate de separare este invariantă prin proiecție, biraportul cu semn $(ABCD)$ este și el invariant. Dacă alegem un punct fix O pe l, ca origine, și dacă alegem drept coordonată x a fiecărui punct de pe l distanța sa orientată, calculată din O, astfel încît coordonatele lui A, B, C, D să fie respectiv x_1, x_2, x_3, x_4, atunci

$$(ABCD) = \frac{CA}{CB} \Big/ \frac{DA}{DB} = \frac{x_3 - x_1}{x_3 - x_2} \Big/ \frac{x_4 - x_1}{x_4 - x_2} = \frac{x_3 - x_1}{x_3 - x_2} \cdot \frac{x_4 - x_2}{x_4 - x_1}.$$

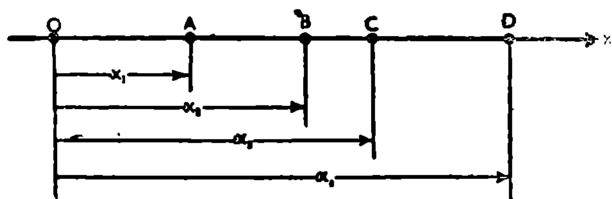


Fig. 78. Biraportul în funcție de coordonate

Dacă $(ABCD) = -1$, astfel încît $CA/CB = -DA/DB$, atunci C și D divid segmentul AB în același raport, unul fiind în interiorul și altul în exteriorul segmentului. În acest caz, se spune că punctele C și D divid armonic segmentul AB, și fiecare din punctele C, D se numește *conjugat armonic* al celuilalt, în raport cu perechea A, B. Dacă $(ABCD) = 1$, atunci punctele C și D (sau A și B) coincid.

Ar trebui să ţinem seama de faptul că *ordinea* în care sînt luate punctele A, B, C, D este o parte esenţială a definiţiei biraportului $(ABCD)$. De exemplu, dacă $(ABCD) = \lambda$, atunci biraportul $(BACD)$ este egal cu $1/\lambda$, în timp ce $(ACBD) = 1 - \lambda$, după cum cititorul poate verifica cu uşurinţă. Patru puncte A, B, C, D pot fi ordonate în $4 \cdot 3 \cdot 2 \cdot 1 = 24$ moduri diferite, fiecare dînd o anumită valoare biraportului lor. Unele dintre aceste permutări vor da aceeaşi valoare biraportului ca şi aranjarea iniţială A, B, C, D; de exemplu, $(ABCD) = (BADC)$. Lăsăm pe seama cititorului, ca exerciţiu, să arate că există doar şase valori diferite ale biraportului pentru aceste 24 permutări diferite ale punctelor, şi anume

$$\lambda, \; 1-\lambda, \; \frac{1}{\lambda}, \; \frac{\lambda-1}{\lambda}, \; \frac{1}{1-\lambda}, \; \frac{\lambda}{\lambda-1}.$$

Aceste şase cantităţi sînt în general distincte, dar două dintre ele pot coincide — ca de pildă, în cazul diviziunii armonice, cînd $\lambda = -1$.

Putem defini, de asemenea *biraportul a patru drepte concurente* şi *coplanare* (adică aflate în acelaşi plan) 1, 2, 3, 4 ca fiind biraportul celor patru puncte de intersecţie ale acestor drepte cu o altă dreaptă, aflată în acelaşi plan. Poziţia celei de-a cincea drepte nu are importanţă, datorită invarianţei biraportului prin proiecţie. Echivalentă cu aceasta este definiţia

$$(1\ 2\ 3\ 4) = \frac{\sin(1,\,3)}{\sin(2,\,3)} \Big/ \frac{\sin(1,\,4)}{\sin(2,\,4)},$$

luată cu semnul plus sau minus, după cum o pereche de drepte nu o separă sau o separă pe cealaltă. (În această formulă, (1, 3), de exemplu, înseamnă unghiul dintre dreptele 1 şi 3.) În sfîrşit, putem defini *biraportul a patru plane coaxiale* (patru plane aflate în spaţiu, care se intersectează după o dreaptă l, axa lor). Dacă o dreaptă intersectează planele în patru puncte, aceste puncte vor avea întotdeauna acelaşi biraport, oricare ar fi poziţia dreptei. (Demonstrarea acestui fapt este lăsată ca exerciţiu.) Deci putem atribui această valoare ca biraport al celor patru plane. În mod echivalent putem defini biraportul a patru plane coaxiale ca fiind biraportul celor patru drepte, după care ele sînt intersectate de un al cincilea plan (fig. 79).

Noţiunea de biraport a patru plane duce natural la întrebarea dacă se poate defini o transformare proiectivă a spaţiului *tridimensional* în el însuşi. Definiţia prin proiecţie centrală nu poate fi generalizată imediat de la două la trei dimensiuni. Însă se poate arăta că orice transformare continuă a unui plan în el însuşi, care leagă într-un mod biunivoc punctele de puncte şi dreptele de drepte, este o transformare proiectivă. Această teoremă sugerează următoarea definiţie pentru trei dimensiuni: O transformare proiectivă a spaţiului este o transformare continuă biunivocă, care conservă dreptele. Se poate arăta că aceste transformări invariază biraportul.

Propozițiile precedente pot fi completate cu cîteva observații. Să presupunem că avem trei puncte diferite A, B, C pe o dreaptă, de coordonate x_1, x_2, x_3. Se cere să se găsească un al patrulea punct D, astfel încît biraportul $(ABCD) = \lambda$, unde λ este dat dinainte. (Cazul particular $\lambda = -1$, pentru care problema se reduce la construirea conjugatului armonic, va fi tratat în mod mai amă-

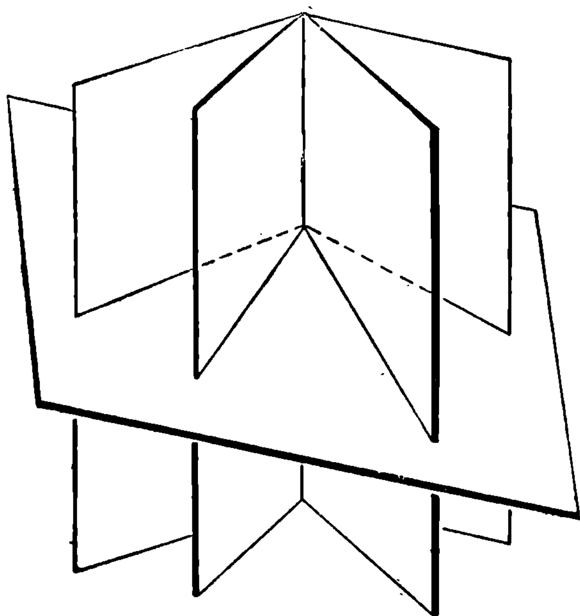


Fig. 79. Biraportul a patru plane coaxiale

nunțit în secțiunea următoare.) În general, problema are o singură soluție; pentru că dacă x este coordonata punctului cerut D, atunci ecuația

$$(2) \qquad \frac{x_3 - x_1}{x_3 - x_2} \cdot \frac{x - x_2}{x - x_1} = \lambda$$

are o singură soluție x. Dacă x_1, x_2, x_3 sînt date și dacă simplificăm ecuația (2), notînd $(x_3 - x_1)/(x_3 - x_2) = k$, prin rezolvarea acestei ecuații găsim că $x = (kx_2 - \lambda x_1)/(k - \lambda)$. De exemplu, dacă cele trei puncte A, B, C sînt echidistante, cu coordonatele $x_1 = 0$, $x_2 = d$, $x_3 = 2d$, atunci $k = (2d - 0)/(2d - d) = 2$ și $x = 2d/(2 - \lambda)$.

Dacă proiectăm o aceeași dreaptă l pe două drepte diferite l', l'' din două centre diferite O' și O'', obținem o corespondență $P \leftrightarrow P'$ între punctele drep

telor l şi l' şi o corespondenţă $P \leftrightarrow P''$ între acelea ale dreptelor l şi l''. Acest fapt stabileşte o corespondenţă $P' \leftrightarrow P''$ între punctele dreptelor l' şi l'', care are proprietatea că orice mulţime de patru puncte A', B', C', D' de pe l' are acelaşi biraport ca şi mulţimea corespunzătoare A'', B'', C'', D'' de pe l''

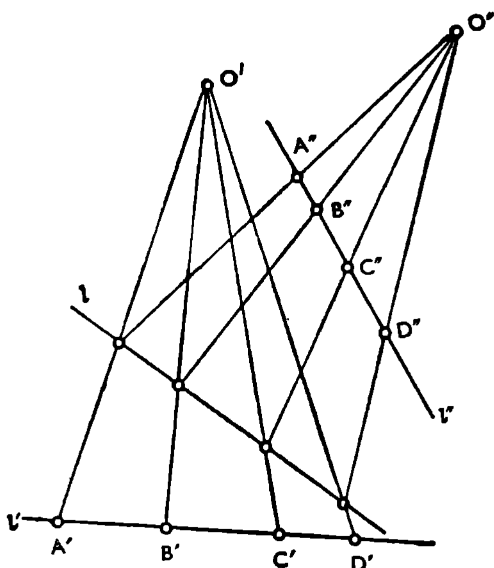


Fig. 80. Corespondenţa proiectivă dintre punctele a două drepte

Orice corespondenţă biunivocă între punctele a două drepte care are această proprietate se numeşte *corespondenţă proiectivă*, indiferent de modul în care ca este definită.

Exerciţii: 1) Arătaţi că, fiind date două drepte şi o corespondenţă proiectivă între punctele lor, putem deplasa o dreaptă printr-o mişcare paralelă într-o astfel de poziţie, încît corespondenţa dată să se obţină printr-o proiecţie simplă. (Indicaţie: suprapuneţi o pereche de puncte corespunzătoare de pe cele două drepte.)

2) Pe baza rezultatului precedent, arătaţi că dacă punctele a două drepte l şi l' sînt puse în corespondenţă printr-o succesiune finită de proiecţii pe diferite drepte intermediare, folosind centre de proiecţie arbitrare, acelaşi rezultat poate fi obţinut doar prin *două* proiecţii.

2. Aplicaţie la patrulaterul complet

Ca aplicaţie interesantă a invarianţei biraportului vom stabili o teoremă simplă, dar importantă, a gemetriei proiective. Ea se referă la *patrulaterul complet*, o figură formată din patru drepte oarecare care nu sînt cîte trei con-

curente, şi din cele şase puncte de intersecţie a lor. În fig. 81, cele patru drepte sînt AE, BE, BI, AF. Dreptele AB, EG şi IF se numesc *diagonalele* patrulaterului. Să luăm o diagonală oarecare, de pildă AB, şi să notăm cu C şi D punctele de intersecţie cu celelalte două diagonale. Atunci avem teorema:

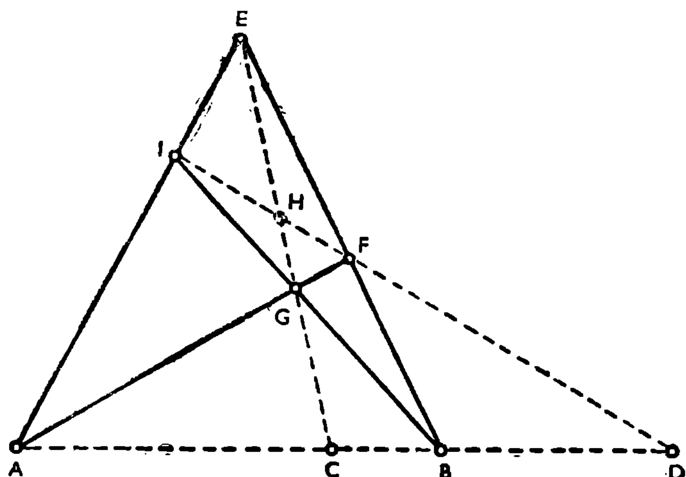


Fig. 81. Patrulaterul complet

$(ABCD) = -1$; cu alte cuvinte, *punctele de intersecţie ale unei diagonale cu celelalte două separă vîrfurile primei diagonale într-o diviziune armonică.* Pentru a demonstra acest lucru, observăm doar că

$x = (ABCD) = (IFHD)$, prin proiecţie din E,

$(IFHD) = (BACD)$, prin proiecţie din G.

Însă ştim că $(BACD) = 1/(ABCD)$; astfel încît $x = 1/x$, $x^2 = 1$, $x = \pm 1$. Deoarece C, D separă pe A, B, biraportul x este negativ şi de aceea trebuie să fie egal cu -1, după cum trebuia demonstrat.

Această proprietate remarcabilă a patrulaterului complet ne permite să găsim numai cu ajutorul riglei conjugatul armonic al unui punct C, în raport cu punctele A, B. Trebuie să alegem doar un punct E, exterior dreptei, să trasăm dreptele EA, EB, EC, să determinăm punctul G pe EC, să trasăm dreapta AG şi dreapta BG, care intersectează dreptele EB şi EA, respectiv în F şi I, iar apoi să trasăm dreapta IF, care intersectează dreapta pe care se află punctele A, B, C în punctul D, conjugat armonic cu C.

Problemă: Dat fiind un segment AB în plan şi o regiune R, după cum se arată în fig. 82, se cere să se prelungească dreapta AB dincolo de R. Cum se poate face acest lucru, doar cu ajutorul riglei, astfel încît rigla să nu intersecteze pe R în timpul construcţiei? (Indicaţie: Alegeţi două puncte arbitrare C, C' pe segmentul AB, apoi determinaţi conjugatele lor armonice D,D' corespunzătoare, cu ajutorul a patru patrulatere, care au punctele A, B ca vîrfuri.)



Fig. 82. Prelungirea unei drepte dincolo de un obstacol

§ 4. PARALELISMUL ŞI INFINITUL

1. Punctele de la infinit ca „puncte ideale"

O examinare a paragrafului precedent va dezvălui faptul că unele raţionamente îşi pierd valabilitatea dacă anumite drepte din construcţii devin paralele. De exemplu, în construcţia de mai sus, al patrulea punct conjugat armonic D nu mai există dacă dreapta IF este paralelă cu AB. Raţionamentul geometric pare să fie împiedicat la fiecare pas de faptul că două drepte paralele nu se intersectează, astfel încît, în orice discuţie care implică intersectarea dreptelor, cazul excepţional al dreptelor paralele trebuie să fie considerat şi formulat în mod separat. De asemenea, proiecţia dintr-un centru O trebuie să fie considerată distinctă de proiecţia paralelă, care necesită o tratare separată. Dacă vrem într-adevăr să discutăm în detaliu fiecare caz excepţional, geometria proiectivă ar deveni foarte complicată. De aceea, trebuie să încercăm o altă cale, şi anume, *să găsim extensiuni ale noţiunilor noastre fundamentale, care să permită eliminarea excepţiilor.*

Aici, intuiţia geometrică ne indică drumul: dacă o dreaptă care intersectează o altă dreaptă este rotită încet spre o poziţie paralelă, atunci punctul de intersecţie al celor două drepte va fugi la infinit. În mod naiv, am putea spune că cele două drepte se intersectează într-un „punct la infinit". Lucrul esenţial este atunci să dăm acestei propoziţii neprecise un înţeles precis, astfel încît punctele de la infinit, sau aşa cum sînt numite uneori, punctele ideale, să poată fi tratate ca şi cum ele ar fi puncte obişnuite din plan sau din spaţiu. Cu alte cuvinte, dorim ca toate regulile referitoare la comportarea punctelor, dreptelor, planelor etc. să rămînă în vigoare, chiar dacă aceste elemente geometrice sînt ideale. Pentru a atinge acest scop, putem proceda fie pe o cale intuitivă, fie pe una formală, tot aşa cum am făcut la extinderea sistemului de numere, unde o

abordare se făcea plecînd de la ideea intuitivă a măsurării, iar alta plecînd de la reguli formale ale operațiilor aritmetice.

Mai întîi trebuie să înțelegem că în geometria sintetică chiar și noțiunile fundamentale „obișnuite" de punct și dreaptă nu sînt definite matematic. Așa-numitele definiții ale acestor noțiuni, care se întîlnesc frecvent în manualele de geometrie elementară, sînt doar descrieri sugestive. În cazul elementelor geometrice obișnuite, intuiția ne dă siguranța „existenței" lor. Însă singurul lucru de care avem într-adevăr nevoie în geometrie, considerată ca sistem matematic, este valabilitatea anumitor reguli, cu ajutorul cărora să putem opera cu aceste noțiuni, ca de pildă unirea punctelor, găsirea intersecției dreptelor etc. Din punct de vedere logic, un „punct" nu este un „lucru în sine", ci este descris în întregime de ansamblul propozițiilor prin care el este legat de alte obiecte. Existența matematică a „punctelor de la infinit" va fi asigurată îndată ce am enunțat limpede și consistent *proprietățile* matematice ale acestor noi entități, adică legăturile lor reciproce și legătura cu punctele „obișnuite". Axiomele obișnuite ale geometriei (de exemplu, cele ale lui Euclid) sînt abstracții ale lumii fizice, ale semnelor făcute cu creionul sau cu creta, ale firelor întinse, ale razelor luminoase, ale barelor rigide etc. Proprietățile pe care aceste axiome le atribuie punctelor și dreptelor matematice sînt descrieri extrem de simplificate și idealizate ale comportării corespondenților lor fizici. Prin două semne făcute cu creionul se pot duce nu una, ci mai multe drepte. Dacă semnele devin din ce în ce mai mici, atunci toate aceste drepte vor avea aproximativ același aspect. Acesta este faptul la care ne gîndim, cînd enunțăm ca axiomă a geometriei faptul că „prin două puncte trece *o dreaptă și numai una*"; nu ne referim la punctele și dreptele fizice, ci la punctele și dreptele abstracte și ideale ale geometriei. Punctele și dreptele geometrice au proprietăți esențial mai simple decît oricare obiecte fizice și această simplificare oferă condiția esențială pentru dezvoltarea geometriei ca știință deductivă.

După cum am remarcat, geometria obișnuită a punctelor și dreptelor se complică în mare măsură, datorită faptului că o pereche de drepte paralele nu se intersectează într-un punct. De aceea, sîntem conduși la o nouă simplificare în structura geometriei, lărgind noțiunea de punct geometric cu scopul de a îndepărta această excepție, tot așa cum am lărgit noțiunea de număr, cu scopul de a îndepărta restricțiile impuse scăderii și împărțirii. Și aici vom fi conduși peste tot de dorința de a conserva, în domeniul extins, legile care guvernau domeniul inițial.

De aceea, vom conveni să adăugăm punctelor obișnuite de pe fiecare dreaptă un singur punct „ideal". Acest punct va fi considerat ca aparținînd tuturor dreptelor paralele cu o dreaptă dată și numai lor. Ca o consecință a acestei convenții, *orice* pereche de drepte din plan se vor intersecta într-un singur punct; dacă dreptele nu sînt paralele, ele se vor intersecta într-un punct obișnuit, în timp ce dacă dreptele sînt paralele, ele se vor intersecta în punctul ideal comun

celor două drepte. Din motive intuitive punctul ideal de pe o dreaptă se numeşte *punctul de la infinit al dreptei.*

Noţiunea intuitivă de punct de pe o dreaptă, care fuge la infinit, ar putea sugera ideea să adăugăm *două* puncte ideale fiecărei drepte, cîte unul corespunzător fiecărui sens. Motivul pentru care adăugăm unul singur, aşa cum am făcut, este că dorim să păstrăm legea după care prin oricare două puncte trece *o singură dreaptă*. Dacă o dreaptă ar conţine două puncte la infinit, comune cu orice dreaptă paralelă, atunci prin aceste două „puncte" ar trece o infinitate de drepte paralele.

Vom conveni, de asemenea, să adăugăm dreptelor obişnuite din plan o singură dreaptă „ideală" (numită, de asemenea, dreapta de la infinit a planului), *care conţine toate punctele ideale ale planului şi numai pe acestea.* Tocmai această convenţie ne este impusă, dacă dorim să păstrăm legea iniţială potrivit căreia prin oricare două puncte se poate duce o dreaptă şi noua lege obţinută, potrivit căreia oricare două drepte au un punct de intersecţie. Pentru a vedea acest lucru, să alegem două puncte ideale oarecare. Atunci singura dreaptă care trebuie să treacă prin aceste puncte nu poate fi o dreaptă obişnuită, deoarece prin convenţia noastră orice dreaptă obişnuită conţine un singur punct ideal. Mai mult, această dreaptă nu poate conţine puncte obişnuite, deoarece un punct obişnuit şi un punct ideal determină o dreaptă obişnuită. În sfîrşit, această dreaptă trebuie să conţină *toate* punctele ideale, dacă vrem ca ea să aibă un punct comun cu orice dreaptă obişnuită. Prin urmare, această dreaptă trebuie să aibă tocmai proprietăţile pe care le-am atribuit dreptei ideale din plan.

Potrivit convenţiilor adoptate, un punct de la infinit este determinat sau este reprezentat de orice familie de drepte paralele, tot aşa cum un număr iraţional este determinat printr-un şir descrescător de intervale raţionale. Afirmaţia că intersecţia a două drepte paralele este un punct de la infinit nu are un subînţeles misterios, ci este doar un mod convenabil de a enunţa faptul că dreptele sînt paralele. Acest mod de a exprima paralelismul, prin limbajul rezervat la început pentru obiecte intuitiv diferite, are ca unic scop să evite considerarea cazurilor excepţionale; ele sînt cuprinse acum în mod automat în acelaşi fel de expresii lingvistice sau de simboluri, care sînt folosite pentru cazurile „obişnuite".

Pentru a rezuma: convenţiile referitoare la punctele de la infinit au fost făcute în aşa fel, încît legile care guvernează relaţia de incidenţă între punctele şi dreptele obişnuite să continue să fie valabile în domeniul extins de puncte, în timp ce operaţia de găsire a punctului de intersecţie a două drepte, iniţial posibilă numai dacă dreptele nu erau paralele, poate fi efectuată acum fără restricţie. Consideraţiile care au dus la această simplificare formală a proprietăţilor relaţiei de incidenţă pot părea oarecum abstracte. Însă ele sînt pe deplin justificate de rezultate, după cum cititorul va vedea în paginile următoare.

2. Elementele ideale și proiecția

Introducerea punctelor de la infinit și a dreptei de la infinit într-un plan ne permit să tratăm proiecția unui plan pe altul într-un mod mult mai satisfăcător. Să considerăm proiecția unui plan π pe un plan π', dintr-un centru

Fig. 83. *Proiecție în elementele de la infinit*

O (fig. 83). Această proiecție stabilește o corespondență între punctele și dreptele lui π și acelea ale lui π'. Fiecărui punct A al lui π îi corespunde un punct unic A' al lui π', cu următoarele excepții: Dacă raza de proiecție dusă prin O este *paralelă* cu planul π', atunci ea intersectează planul π într-un punct A, căruia nu-i corespunde nici un punct obișnuite al lui π'. Aceste puncte excepționale ale lui π se află pe o dreaptă l, căreia nu-i corespunde nici o dreaptă obișnuită a lui π'. Însă aceste excepții sînt eliminate, dacă facem convenția că lui A îi corespunde punctul de la infinit din π', aflat în direcția dreptei OA și că lui l îi corespunde dreapta de la infinit din π'. În același mod asociem un punct de la infinit în π oricărui punct B' de pe dreapta m' din π', prin care trec toate razele duse prin O paralele cu planul π. Lui m' îi va corespunde dreapta de la infinit din π. Astfel, prin introducerea punctelor și dreptelor de la infinit într-un plan, *proiecția unui plan pe un altul stabilește o corespondență între punctele și dreptele celor două plane, care este biunivocă fără excepție*. (Aceasta ne scutește de excepțiile menționate în nota de la p.188). Mai mult, se vede cu ușurință că o consecință a convenției noastre este faptul că un *punct se află pe o dreaptă*,

dacă și numai dacă proiecția punctului se află pe proiecția dreptei. Prin urmare, toate propozițiile referitoare la puncte coliniare, drepte concurente etc., care se referă doar la puncte, drepte și la relații de incidență, se arată a fi invariante prin proiecție în sensul extins. Aceasta ne permite să operăm cu punctele de la infinit dintr-un plan π, operînd doar cu punctele obișnuite care le corespund, într-un plan π', legat de π printr-o proiecție.

*Interpretarea punctelor de la infinit ale unui plan π cu ajutorul proiecției dintr-un punct exterior O, pe punctele obișnuite dintr-un alt plan π', poate fi folosită pentru a da un „model" euclidian concret al planului extins. În acest scop, neglijăm planul π', și ne fixăm atenția asupra planului π și a dreptelor duse prin O. Oricărui punct obișnuit al lui π îi corespunde o dreaptă dusă prin O neparalelă cu π; oricărui punct de la infinit al lui π îi corespunde o dreaptă dusă prin O, paralelă cu π. Deci totalității punctelor obișnuite și ideale ale lui π îi corespunde totalitatea dreptelor duse prin O, iar această corespondență este biunivocă, fără excepție. *Punctele* de pe o *dreaptă* a lui π vor corespunde *dreptelor* dintr-un *plan* care trec prin O. Un punct și o dreaptă din π vor fi incidente, dacă și numai dacă dreapta și planul corespunzător, duse prin O, sînt incidente. Deci geometria incidenței punctelor și dreptelor din planul extins este întru totul echivalentă cu geometria incidenței dreptelor și planelor obișnuite, care trec printr-un punct fix din spațiu.

* În trei dimensiuni, situația este asemănătoare, cu toate că nu mai putem lămuri lucrurile în mod intuitiv prin proiecție. Introducem din nou un punct la infinit, asociat fiecărei familii de drepte paralele. În fiecare plan avem o dreaptă la infinit. Apoi trebuie să introducem un element nou, *planul de la infinit*, care constă din toate punctele de la infinit ale spațiului și care conține toate dreptele de la infinit. Orice plan obișnuit intersectează planul de la infinit după dreapta sa de la infinit.

3. Biraportul cu elementele de la infinit

Trebuie să facem o observație referitoare la biraportul format cu elemente de la infinit. Să notăm punctul de la infinit al unei drepte l prin simbolul ∞. Dacă A, B, C sînt trei puncte obișnuite ale lui l, atunci putem atribui o valoare simbolului $(ABC\infty)$ în modul următor: să alegem un punct P pe l; atunci $(ABC\infty)$ trebuie să fie limita către care tinde $(ABCP)$, cînd P tinde la infinit, rămînînd pe l. Dar

$$(ABCP) = \frac{CA}{CB} \Big/ \frac{PA}{PB},$$

și cînd P tinde la infinit, PA/PB tinde către 1. Deci definim

$$(ABC\infty) = CA/CB.$$

În particular dacă $(ABC\infty) = -1$, atunci C este mijlocul segmentului AB: *mijlocul și punctul de la infinit, luat în direcția segmentului, împart armonic segmentul dat.*

Exerciții: Care este biraportul a patru drepte l_1, l_2, l_3, l_4, dacă ele sînt paralele? Care este biraportul lor dacă l_1 este dreapta de la infinit?

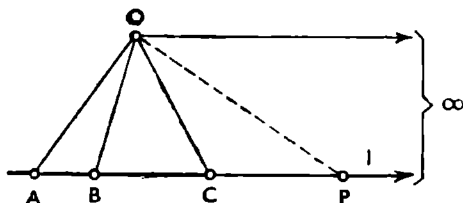


Fig. 84. Biraportul cu un punct la infinit

§ 5. APLICAȚII

1. Observații preliminare

Introducînd elementele de la infinit nu mai este necesar să enunțăm explicit cazurile excepționale care apar în construcții și teoreme, cînd două sau mai multe drepte sînt paralele. Trebuie doar să ne reamintim că dacă un punct este la infinit, toate dreptele care trec prin acel punct sînt paralele. Distincția făcută între proiecția centrală și cea paralelă nu mai este necesară, deoarece aceasta din urmă înseamnă, pur și simplu, proiecția dintr-un punct de la infinit. În fig. 72, punctul O sau dreapta PQR poate fi la infinit (fig. 85 arată primul caz); lăsăm ca exercițiu pe seama cititorului să formuleze în limbaj „finit" enunțul corespunzător al teoremei lui Desargues.

Nu numai *enunțul*, dar și *demonstrația* unei teoreme proiective este adesea mai simplă dacă utilizăm elementele de la infinit. Principiul general este următorul. Prin „clasa proiectivă" a unei figuri geometrice F înțelegem clasa tuturor figurilor în care figura F poate fi transformată prin transformări proiective. Proprietățile proiective ale lui F vor fi identice cu acelea ale oricărui membru al clasei sale proiective, deoarece, prin definiție, proprietățile proiective sînt invariante prin proiecție. Astfel, orice teoremă proiectivă (care se referă numai la proprietăți proiective) care este adevărată pentru F va fi adevărată pentru orice membru al clasei proiective a lui F și reciproc. Prin urmare, pentru a demonstra o astfel de teoremă pentru F, este suficient să o demonstrăm pentru oricare alt membru al clasei proiective a lui F. Putem folosi adesea acest principiu, găsind un membru particular al clasei proiective a lui F, pentru care teorema este mai ușor de demonstrat decît pentru F însăși. De exemplu, orice două puncte A, B ale planului π pot fi proiectate la infinit,

proiectîndu-le dintr-un centru O pe un plan π', paralel cu planul care trece prin O, A, B; dreptele care trec prin A şi acelea care trec prin B vor fi transformate în două familii de drepte paralele. În teoremele proiective, pe care le vom demonstra în acest paragraf, vom face o astfel de transformare preliminară.

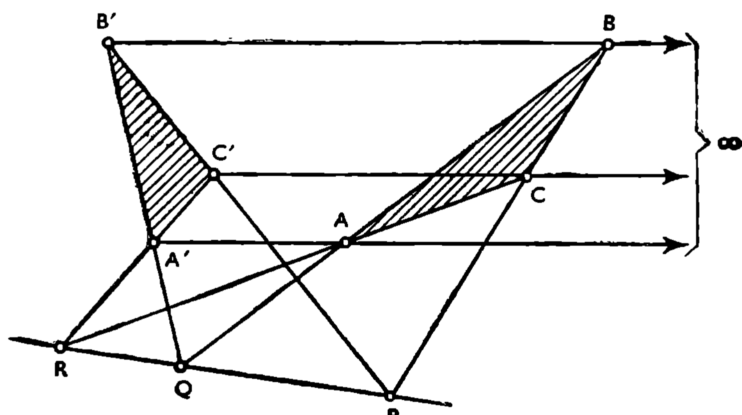


Fig. 85. Configurația lui Desargues cu centrul la infinit

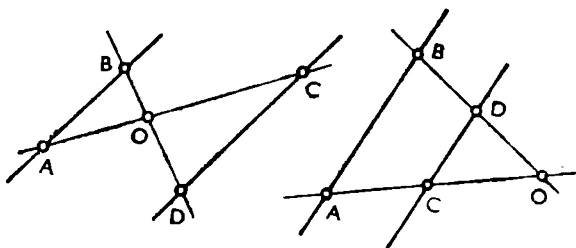


Fig. 86.

Următorul fapt elementar referitor la drepte paralele va fi util. Să presupunem că două drepte, care se intersectează în punctul O, sînt intersectate de două drepte l_1 şi l_2 în punctele A, B, C, D, după cum se arată în fig. 86. Dacă l_1 şi l_2 sînt paralele, atunci

$$\frac{OA}{OC}=\frac{OB}{OD};$$

şi reciproc, dacă $OA/OC = OB/OD$, atunci l_1 şi l_2 sînt paralele. Demonstraţia rezultă din proprietăţi elementare ale triunghiurilor asemenea şi va fi lăsată pe seama cititorului.

2. Demonstraţia teoremei lui Desargues în plan

Demonstrăm acum că dacă pentru două triunghiuri ABC şi $A'B'C'$ dintr-un plan, situate aşa cum se arată în fig. 72, dreptele care trec prin vîrfurile corespunzătoare se intersectează într-un punct, intersecţiile P, Q, R ale laturilor corespunzătoare se află pe o dreaptă. Pentru a face acest lucru, proiectăm mai întîi figura, astfel încît Q şi R să meargă la infinit. După proiecţie, AB va fi paralelă cu $A'B'$, AC cu $A'C'$ şi figura va arăta ca în fig. 87. După cum am indicat în secţiunea 1 a acestui paragraf, pentru a demonstra teorema lui Desargues în general, este suficient să o demonstrăm pentru acest caz particular de figură. În acest scop, trebuie să arătăm doar că intersecţia lui BC cu $B'C'$ se află, de asemenea, la infinit, astfel încît BC este paralelă cu $B'C'$; atunci P, Q, R vor fi într-adevăr coliniare. (Pentru că ele se vor afla pe dreapta de la infinit.) Însă

$$AB \,||\, A'B' \text{ implică } \frac{u}{v} = \frac{r}{s},$$

şi

$$AC \,||\, A'C' \text{ implică } \frac{x}{y} = \frac{r}{s}.$$

De aceea, $u/v = x/y$; aceasta implică $BC \,||\, B'C'$, ceea ce trebuia demonstrat.

Remarcaţi că această demonstraţie a teoremei lui Desargues foloseşte noţiunea metrică de lungime a unui segment. Astfel, am demonstrat o teoremă

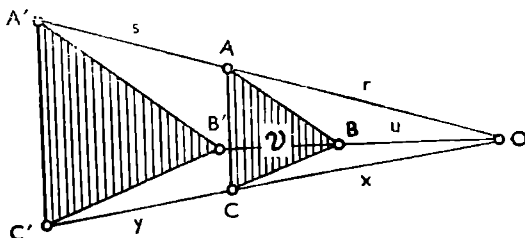


Fig. 87. Demonstrarea teoremei lui Desargues

proiectivă prin metode metrice. Mai mult, dacă transformările proiective sînt definite „intrinsec" ca transformări plane care conservă biraportul (cf. p.195), atunci demonstraţia rămîne în întregime în plan.

Exercițiu: Demonstrați în mod similar reciproca teoremei lui Desargues: Dacă triunghiurile ABC și $A'B'C'$ au proprietatea că punctele P, Q, R sînt coliniare, atunci dreptele AA', BB', CC' sînt concurente.

3. Teorema lui Pascal[6]

Această teoremă are următorul enunț: *Dacă vîrfurile unui hexagon se află alternativ pe o pereche de drepte concurente, atunci cele trei intersecții* P, Q, R *ale laturilor opuse ale hexagonului sînt coliniare* (fig. 88). (Hexagonul se poate intersecta pe sine însuși. Laturile „opuse" pot fi recunoscute din diagrama schematică a fig. 89.)

Efectuînd o proiecție preliminară, putem presupune că P și Q se află la infinit. Atunci trebuie să arătăm doar că R este și el la infinit. Situația este ilustrată în fig. 90, unde 23 || 56 și 12 || 45. Trebuie să arătăm că 16 || 34. Avem

$$\frac{a}{a+x}=\frac{b+y}{b+y+s}, \quad \frac{b}{b+y}=\frac{a+x}{a+x+r}.$$

De aceea

$$\frac{a}{b}=\frac{a+x+r}{b+y+s},$$

astfel încît 16 || 34, după cum trebuia demonstrat.

4. Teorema lui Brianchon

Această teoremă afirmă următoarele: *Dacă laturile unui hexagon trec alternativ prin două puncte fixe,* P *și* Q *atunci cele trei diagonale care unesc perechile opuse de vîrfuri ale hexagonului sînt concurente* (fig. 91). Printr-o proiecție putem trimite la infinit punctul P și punctul în care două dintre diagonale, de pildă 14 și 36, se intersectează. Atunci situația va apare ca în fig. 92. Deoarece 14 || 36, avem $a/b = u/v$. Însă $x/y = a/b$ și $u/v = r/s$. De aceea, $x/y = r/s$ și 36 || 25, astfel încît toate cele trei diagonale sînt paralele și deci concurente. Aceasta este suficient pentru a demonstra teorema în cazul general.

[6] La p. 228 vom discuta o teoremă mai generală de același tip. Cazul particular tratat aici mai este cunoscut și sub numele descoperitorului său, Pappus din Alexandria (secolul al III-lea e.n.).

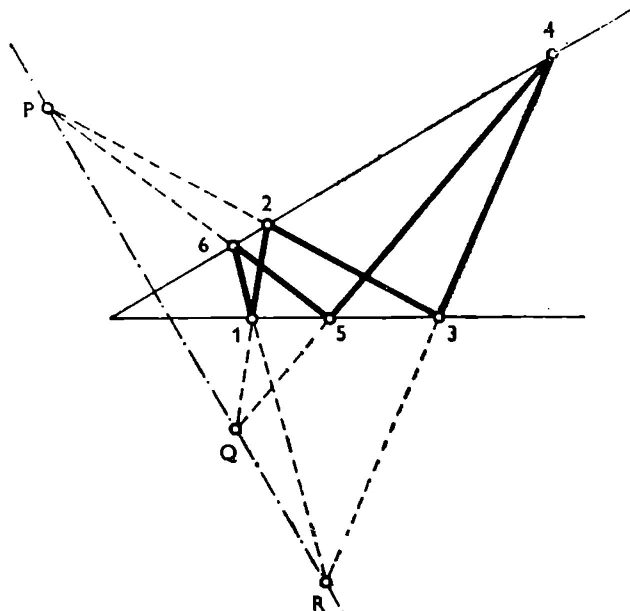


Fig. 88. Configurația lui Pascal

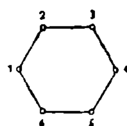


Fig. 89.

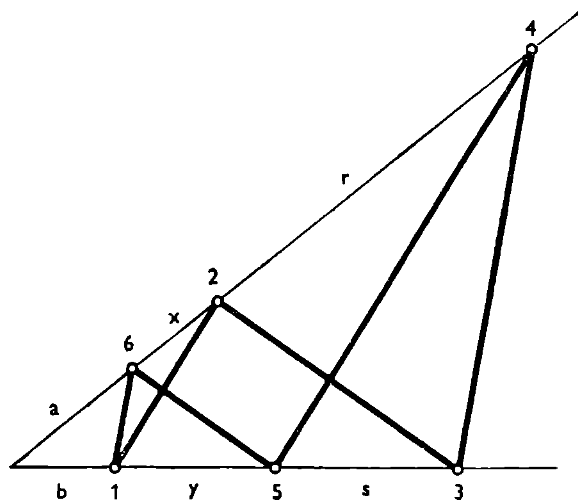


Fig. 90. Demonstrarea teoremei lui Pascal

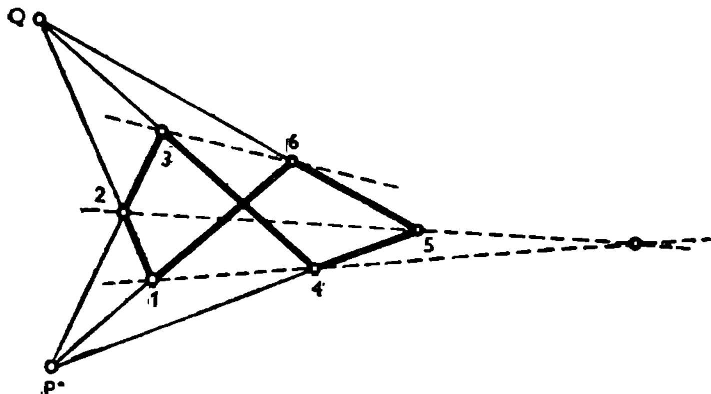


Fig. 91. Configurația lui Brianchon

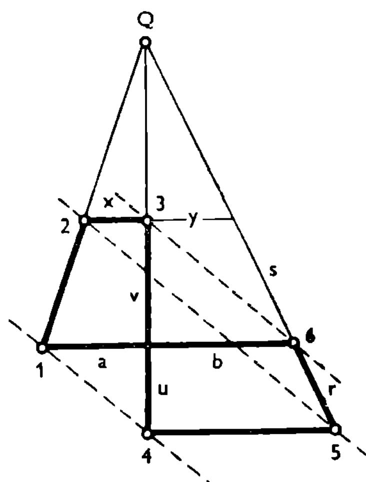


Fig. 92. Demonstrarea teoremei lui Brianchon

5. Observație asupra dualității

Cititorul trebuie să fi observat asemănarea remarcabilă dintre teoremele lui Pascal (1623–1662) și Brianchon (1785–1864). Această asemănare devine deosebit de impresionantă, dacă scriem teoremele alăturat.

Teorema lui Pascal	*Teorema lui Brianchon*
Dacă *vîrfurile* unui hexagon *se află alternativ pe* două *drepte*, atunci *punctele de intersecție ale laturilor opuse* sînt *coliniare*.	Dacă *laturile* unui hexagon *trec alternativ prin* două *puncte*, atunci *dreptele care unesc vîrfurile opuse* sînt *concurente*.

Nu numai teoremele lui Pascal și Brianchon, dar toate teoremele geometriei proiective se grupează în perechi de teoreme asemănătoare și de structură oarecum identică. Această legătură se numește *dualitate*. În geometria plană, punctul și dreapta se numesc *elemente duale*. Trasarea unei drepte printr-un punct și alegerea unui punct pe o dreaptă sînt *operații duale*. Două figuri sînt duale, dacă una poate fi obținută din cealaltă, înlocuind fiecare element și operație prin elementul sau operația duală. Două teoreme sînt duale, dacă una se transformă în alta, atunci cînd toate elementele și operațiile sînt înlocuite prin dualele lor. De exemplu, teoremele lui Pascal și Brianchon sînt duale, iar duala teoremei lui Desargues este chiar reciproca ei. Acest fenomen de dualitate conferă geometriei proiective un caracter cu totul distinct de acela al geometriei elementare (metrice), în care nu există o astfel de dualitate (de exemplu, ar fi fără sens să vorbim despre dualul unui unghi de 37° sau al unui segment de lungime 2). În multe manuale de geometrie proiectivă se indică *principiul dualității*, care afirmă că *duala oricărei teoreme adevărate a geometriei proiective este de asemenea o teoremă adevărată a geometriei proiective*, așezînd teoremele duale împreună cu demonstrațiile lor duale pe coloane paralele ale paginii, așa cum am făcut mai sus. Motivul fundamental al acestei dualități va fi considerat în paragraful următor (a se vedea, de asemenea, p. 235).

§ 6. REPREZENTAREA ANALITICĂ

1. Observații introductive

În prima etapă a dezvoltării geometriei proiective, exista o tendință puternică de a construi totul pe o bază sintetică și „pur geometrică", evitînd folosirea numerelor și a metodelor algebrice. Acest program a întîmpinat mari dificultăți, deoarece întotdeauna rămîneau locuri în care o anumită formulare algebrică părea inevitabilă. Un succes deplin în construirea unei geometrii proiective pur sintetice a fost obținut doar spre sfîrșitul secolului al XIX-lea, cu prețul unei complicări destul de mari. În această privință, metodele geometriei analitice au avut mai mult succes. Tendința generală în matematica modernă este de a baza totul pe noțiunea de număr, iar în geometrie, această tendință care a început cu Fermat și Descartes, a avut succese hotărîtoare. Geometria analitică s-a dezvoltat din stadiul unui simplu instrument în ratio-

namentul geometric, într-un domeniu de sine stătător în care interpretarea geometrică intuitivă a operațiilor și a rezultatelor nu mai este scopul ultim și exclusiv, ci are mai degrabă rolul unui principiu călăuzitor, care ajută la sugerarea și înțelegerea rezultatelor analitice. Schimbarea sensului geometriei este rezultatul unei dezvoltări istorice treptate, care a lărgit în mod considerabil cîmpul geometriei clasice și în același timp a realizat o sinteză aproape organică a geometriei și analizei.

În geometria analitică, „coordonatele" unui obiect geometric sînt orice sistem de numere care caracterizează obiectul în mod unic. Astfel, un punct este definit prin indicarea coordonatelor sale rectangulare x, y sau a coordonatelor polare ρ, θ, în timp ce un triunghi poate fi definit prin indicarea coordonatelor celor trei vîrfuri, ceea ce necesită în total șase coordonate. Știm că o dreaptă din planul x, y este locul geometric al punctelor $P(x, y)$ (pentru notație cf. p. 91), ale căror coordonate satisfac o ecuație liniară de forma

$$ax + by + c = 0. \tag{1}$$

De aceea, putem numi cele trei numere a, b, c, „coordonatele" acestei drepte. De exemplu, $a = 0$, $b = 1$, $c = 0$ definesc dreapta $y = 0$, care este axa Ox; $a = 1$, $b = -1$, $c = 0$ definesc dreapta $x = y$, care este prima bisectoare a sistemului de coordonate. În același mod, ecuațiile pătratice definesc „secțiuni conice":

$x^2 + y^2 = r^2$ un cerc cu centrul în origine de rază r,

$(x - a)^2 + (y - b)^2 = r^2$ un cerc cu centrul în (a, b) de rază r,

$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ o elipsă.

etc.

Abordarea naivă a geometriei analitice se face pornind de la noțiuni pur „geometrice" — punct, dreaptă etc. — care sînt traduse apoi în limbajul numerelor. Punctul de vedere modern este opus. Începem cu *mulțimea tuturor perechilor de numere* x, y și *numim* punct orice astfel de pereche, deoarece dacă vrem, putem *interpreta* intuitiv o astfel de pereche de numere prin noțiunea familiară de punct geometric. În mod asemănător, o ecuație liniară între x și y definește o dreaptă. O astfel de deplasare a accentului de la aspectul intuitiv al geometriei spre cel analitic deschide drumul unei tratări simple, dar riguroase, a punctelor de la infinit din geometria proiectivă și este indispensabilă pentru o înțelegere mai profundă a întregului subiect. Pentru cititorii care posedă o anumită formație preliminară vom da o expunere a acestui mod de abordare.

*2. Coordonate omogene. Baza algebrică a dualității

În geometria analitică obișnuită, coordonatele rectangulare ale unui punct din plan sînt distanțe orientate ale punctului la o pereche de axe perpendiculare. Acest sistem își pierde valabilitatea pentru punctele de la infinit din planul extins al geometriei proiective. Deci, dacă dorim să aplicăm metodele analitice geometriei proiective, este necesar să găsim un sistem de coordonate, care să cuprindă atît punctele ideale, cît și pe cele obișnuite. Introducerea unui astfel de sistem de coordonate este descrisă cel mai bine, făcînd ipoteza că planul X, Y, pe care să-l notăm cu π, este scufundat în spațiul tridimensional, în care au fost introduse coordonatele rectangulare x, y, z (distanțele orientate ale unui punct la trei plane coordonate, determinate de axele Ox, Oy, Oz). Așezăm planul π paralel cu planul xOy, la o distanță egală cu 1, deasupra lui, astfel încît orice punct P al lui π va avea coordonatele tridimensionale $(X, Y, 1)$. Luînd originea O a sistemului de coordonate drept centru de proiecție, observăm că *orice punct* P *determină o singură dreaptă care trece prin* O *și reciproc* (cf. p. 202. Dreptele care trec prin O și sînt paralele cu π corespund punctelor de la infinit ale lui π).

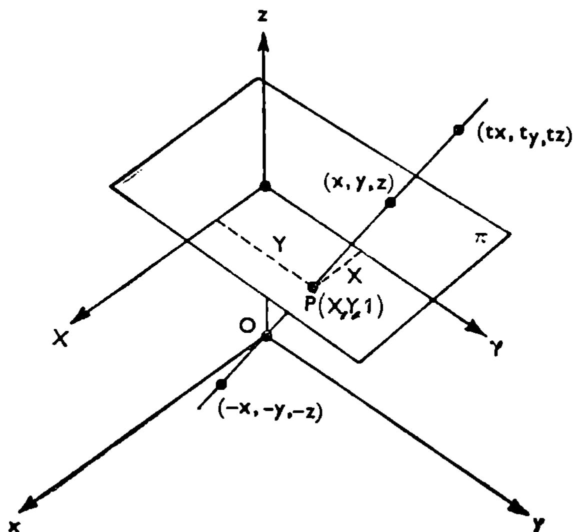


Fig. 93. Coordonate omogene

Vom descrie acum un sistem de „coordonate omogene" pentru punctele lui π. Pentru a găsi coordonatele omogene ale oricărui punct obișnuit P al lui π, luăm dreapta care trece prin O și P și alegem un punct *oarecare* Q, diferit de O, de pe această dreaptă. (fig. 93). Atunci coordonatele tridimensionate obișnuite

x, y, z ale lui Q se numesc *coordonatele omogene* ale lui P. În particular, coordonatele $(X, Y, 1)$ ale lui P sînt coordonate omogene ale lui P. Mai mult, orice mulțime de numere (tX, tY, t), cu $t \neq 0$, va fi de asemenea un sistem de coordonate omogene ale lui P, deoarece coordonatele tuturor punctelor de pe dreapta OP, diferite de O, vor fi de această formă. (Am exclus punctul $(0, 0, 0)$, deoarece el se află pe toate dreptele care trec prin O și nu poate face nici o distincție între ele.)

Această metodă de a introduce coordonatele în plan necesită cunoașterea a trei numere în loc de două, pentru a preciza poziția unui punct, și mai are neajunsul că coordonatele unui punct nu sînt determinate în mod unic, ci doar făcînd abstracție de un factor $t \neq 0$. Însă ea prezintă marele avantaj că punctele de la infinit ale lui π sînt incluse acum în reprezentarea cu ajutorul coordonatelor. Un punct P de la infinit din planul π este determinat de o dreaptă care trece prin O, paralelă cu π. Orice punct Q de pe această dreaptă va avea coordonatele de forma $(x, y, 0)$. Deci *coordonatele omogene ale unui punct de la infinit din* π *sînt de forma* $(x, y, 0)$.

Ecuația unei drepte din π în coordonate omogene poate fi găsită imediat, observînd că dreptele care unesc punctul O cu punctele acestei drepte se află într-un plan care trece prin O. Se arată în geometria analitică că ecuația unui astfel de plan este de forma

$$ax + by + cz = 0.$$

Deci aceasta este ecuația unei drepte din π, în coordonate omogene.

Acum, după ce modelul geometric al punctelor lui π, reprezentate prin drepte care trec prin O, și-a îndeplinit scopul, îl putem lăsa la o parte și dăm următoarea definiție pur analitică a planului extins:

Un *punct* este un triplet ordonat de numere reale (x, y, z), nu toate nule. Două astfel de triplete (x_1, y_1, z_1) și (x_2, y_2, z_2) definesc *același* punct dacă există un $t \neq 0$, astfel încît

$$x_2 = tx_1, \quad y_2 = ty_1, \quad z_2 = tz_1.$$

Cu alte cuvinte, coordonatele oricărui punct pot fi înmulțite cu orice factor nenul, fără ca prin aceasta să schimbăm definiția punctului. (Din acest motiv ele se numesc coordonate *omogene*.) Un punct (x, y, z) este un punct *obișnuit*, dacă $z \neq 0$; dacă $z = 0$, el este un *punct de la infinit*.

O *dreaptă* din π este formată din toate punctele (x, y, z) care satisfac o ecuație liniară de forma

$$ax + by + cz = 0, \tag{1'}$$

unde a, b, c sînt trei constante, nu toate nule. În particular, toate punctele de la infinit din π satisfac ecuația liniară

$$z = 0. \tag{2}$$

Aceasta este prin definiţie o dreaptă şi se numeşte *dreapta de la infinit* din π. Deoarece o dreaptă este definită printr-o ecuaţie de forma (1'), numim tripletul de numere (a, b, c) *coordonatele omogene ale dreptei* (1'). Rezultă că (ta, tb, tc) pentru orice $t \neq 0$, sînt de asemenea coordonatele dreptei (1'), deoarece ecuaţia

$$(ta)x + (tb)y + (tc)z = 0 \tag{3}$$

este satisfăcută de aceleaşi triplete de coordonate (x, y, z), ca şi ecuaţia (1').

În aceste definiţii observăm simetria perfectă dintre punct şi dreaptă: fiecare este precizat prin trei coordonate omogene (u, v, w). Condiţia ca punctu- (x, y, z) să se afle pe dreapta (a, b, c) este ca

$$ax + by + cz = 0,$$

şi aceasta este, de asemenea, condiţia ca punctul ale cărui coordonate sînt (a, b, c) să se afle pe dreapta ale cărei coordonate sînt (x, y, z). De exemplu, identitatea aritmetică

$$2 \cdot 3 + 1 \cdot 4 - 5 \cdot 2 = 0$$

poate fi interpretată atît prin faptul că punctul (3, 4, 2) se află pe dreapta (2, 1, —5), cît şi prin faptul că punctul (2, 1, —5) se află pe dreapta (3, 4, 2) Această simetrie este baza dualităţii între punct şi dreaptă în geometria proiectivă, pentru că orice relaţie între puncte şi drepte devine o relaţie între drepte şi puncte, dacă coordonatele sînt reinterpretate în mod potrivit. În noua interpretare, coordonatele iniţiale ale punctelor şi dreptelor sînt gîndite acum ca reprezentînd respectiv drepte şi puncte. Toate operaţiile algebrice şi rezultatele rămîn neschimbate, însă interpretarea lor dă imaginea duală a teoremei iniţiale. Trebuie să observăm că această dualitate nu se menţine în planul obişnuit al celor două coordonate X, Y, deoarece ecuaţia unei drepte în coordonate obişnuite

$$aX + bY + c = 0$$

nu este simetrică în raport cu X, Y pe de o parte şi cu a, b, c pe de alta. Numai prin includerea punctelor şi a dreptei de la infinit putem stabili principiul dualităţii.

Pentru a trece de la coordonatele omogene x, y, z ale unui punct obişnuit P din planul π la coordonate rectangulare obişnuite, punem, pur şi simplu, $X = x/z$, $Y = y/z$. Atunci X, Y reprezintă distanţele de la punctul P la două axe perpendiculare, axa absciselor şi axa ordonatelor din planul π, după cum se arată în fig. 93. Ştim că o ecuaţie de forma

$$aX + bY + c = 0$$

reprezintă o dreaptă din π. Făcînd substituţia $X = x/z$, $Y = y/z$ şi înmulţind totul cu z, găsim că ecuaţia aceleiaşi drepte în coordonate omogene este

$$ax + by + cz = 0,$$

aşa cum s-a afirmat la p. 214. Astfel, ecuaţia dreptei $2x - 3y + z = 0$ în coordonate rectangulare X, Y devine $2X - 3Y + 1 = 0$. Desigur, ultima ecuaţie nu este verificată de punctul de la infinit al acestei drepte, care are drept coordonate omogene tripletul (3, 2, 0).

Mai rămîne de spus încă ceva. Am reuşit să dăm o definiţie pur analitică a punctului şi a dreptei, dar ce putem spune despre noţiunea importantă de transformare proiectivă? Se poate arăta că o transformare proiectivă a unui plan pe un altul, aşa cum a fost definită la p. 188, este dată analitic de un sistem de ecuaţii liniare

$$
\begin{aligned}
x' &= a_1x + b_1y + c_1z,\\
y' &= a_2x + b_2y + c_2z, \qquad (4)\\
z' &= a_3x + b_3y + c_3z,
\end{aligned}
$$

care leagă coordonatele omogene x', y', z' ale punctelor din planul π', de coordonatele omogene x, y, z ale punctelor din planul π. Din punctul nostru de vedere, putem *defini* acum o transformare proiectivă ca fiind transformarea dată de orice sistem de ecuaţii liniare de forma (4). Teoremele geometriei proiective devin atunci teoreme referitoare la comportarea tripletelor de numere (x, y, z) faţă de astfel de transformări. De exemplu, demonstrarea faptului că biraportul a patru puncte de pe o dreaptă rămîne neschimbat prin astfel de transformări devine un exerciţiu simplu din algebra transformărilor liniare. Nu putem intra mai mult în amănuntele procedeului analitic. În schimb, ne vom întoarce la aspectele mai intuitive ale geometriei proiective.

§ 7. PROBLEME DE CONSTRUCŢIE NUMAI CU AJUTORUL RIGLEI

În construcţiile următoare se subînţelege că singurul instrument folosit este rigla.

Problemele 1—18 sînt conţinute într-o lucrare a lui J. Steiner, în care el demonstrează că pentru construcţiile geometrice ne putem lipsi de compas, dacă este dat un cerc fixat, împreună cu centrul său (cf. cap. III, p. 168). Cititorul este sfătuit să rezolve aceste probleme în ordinea indicată.

O mulţime formată din patru drepte *a*, *b*, *c*, *d*, care trec printr-un punct *P* se numeşte *armonică* dacă biraportul (*abcd*) este egal cu —1. Dreptele *a* şi *b* se numesc *conjugate* în raport cu *c* şi *d*, şi reciproc.

1) Demonstraţi că: dacă într-un sistem format din patru drepte armonice *a*, *b*, *c*, *d*, raza *a* este bisectoarea unghiului format de *c* şi *d*, atunci *b* este perpendiculară pe *a*.

2) Construiţi a patra dreaptă armonică faţă de trei drepte date, care trec printr-un punct. (Indicaţie: folosiţi teorema patrulaterului complet.)

3) Construiţi al patrulea punct armonic conjugat faţă de trei puncte date pe o dreaptă.

4) Fiind dat un unghi drept şi un unghi arbitrar, care au vîrfurile şi o latură comună, dublaţi unghiul arbitrar dat.

5) Fiind dat un unghi şi bisectoarea lui *b*, construiţi o perpendiculară pe *b*, în vîrful unghiului dat.

6) Demonstrați că dacă dreptele $l_1, l_2, l_3, \ldots, l_n$, care trec printr-un punct P, intersectează dreapta a în punctele $A_1, A_2, \ldots, A_n$ și intersectează dreapta b în punctele $B_1, B_2, \ldots, B_n$ atunci toate intersecțiile perechilor de drepte A_iB_k și A_kB_i ($i \neq k$; $i, k = 1, 2, \ldots, n$) se află pe o dreaptă.

7) Demonstrați că dacă o paralelă la latura BC a triunghiului ABC intersectează pe AB în B' și pe AC în C', atunci dreapta care unește pe A cu intersecția D a lui $B'C$ și $C'B$ înjumătățește segmentul BC.

7a) Formulați și demonstrați reciproca propoziției din 7.

8) Pe o dreaptă l sînt date trei puncte P, Q, R, astfel încît Q este mijlocul segmentului PR. Construiți o paralelă la l, printr-un punct dat S.

9) Fiind date două drepte paralele l_1 și l_2, înjumătățiți un segment dat AB, de pe l_1.

10) Duceți o paralelă printr-un punct dat P la două drepte paralele date l_1 și l_2. (Indicație: reduceți problema 9 la problema 7, folosind problema 8.)

11) Steiner dă următoarea soluție pentru problema dublării unui segment dat AB, atunci cînd este dată o paralelă l la dreapta AB: printr-un punct C, care nu se află pe l și nici pe dreapta AB, duceți dreapta CA, care intersectează pe l în A_1, și dreapta CB, care intersectează pe l în B_1. Atunci (cf. problema 10) duceți o paralelă la l prin punctul C, care întîlnește pe BA_1 în D. Dacă DB_1 întîlnește pe AB în E, atunci $AE = 2AB$.

Demonstrați ultima propoziție.

12) Împărțiți un segment AB în n părți egale, dată fiind o paralelă l la dreapta AB. (Indicație: construiți mai întîi un segment de n ori mai mare decît un segment arbitrar de pe l, folosind problema 11.)

13) Fiind dat un paralelogram $ABCD$, duceți o paralelă printr-un punct P la o dreaptă dată l. (Indicație: aplicați problema 10 centrului paralelogramului și folosiți problema 8.)

14) Fiind dat un paralelogram, înmulțiți cu n un segment dat. (Indicație: folosiți problemele 13 și 11.)

15) Fiind dat un paralelogram, împărțiți un segment dat în n părți egale.

16) Fiind dat un cerc și centrul său, duceți o paralelă la o dreaptă dată printr-un punct dat. (Indicație: folosiți problema 13.)

17) Fiind dat un cerc și centrul său, înmulțiți și împărțiți cu n un segment dat. (Indicație: folosiți problema 13.)

18) Fiind dat un cerc și centrul său, coborîți o perpendiculară dintr-un punct dat pe o dreaptă dată. (Folosind un dreptunghi înscris în cercul dat, cu două din laturi paralele cu dreapta dată, reduceți problema la exercițiul precedent.)

19) Care sînt problemele fundamentale de construcție pe care le puteți rezolva cu ajutorul rezultatelor problemelor 1—18, dacă instrumentul dv. este o riglă cu două muchii paralele?

20) Două drepte l_1 și l_2 date, se intersectează într-un punct P, *exterior* foii pe care sînt desenate. Construiți dreapta care unește un punct dat Q cu P. (Indicație: completați elementele date pînă la configurația figurii din teorema plană a lui Desargues, astfel încît P și Q să devină intersecțiile laturilor corespunzătoare a două triunghiuri din teorema lui Desargues.)

21) Construiți dreapta care unește două puncte date, aflate la o distanță mai mare decît lungimea riglei folosite. (Indicație: folosiți problema 20.)

22) Două puncte P şi Q, exterioare foii de hîrtie, sînt determinate de două perechi de drepte l_1, l_2 şi m_1, m_2, care trec respectiv prin P şi Q. Construiţi acea parte a dreptei PQ, care se află pe foaia de hîrtie dată. (Indicaţie: pentru a obţine un punct al dreptei PQ, completaţi elementele date pînă la configuraţia teoremei lui Desargues, astfel încît unul din triunghiuri să aibă două laturi pe l_1 şi m_1, iar celălalt să aibă laturile corespunzătoare pe l_2 şi m_2.)

23) Rezolvaţi problema 20 cu ajutorul teoremei lui Pascal. (p. 207). (Indicaţie: completaţi elementele date pînă la configuraţia teoremei lui Pascal, folosind dreptele l_1, l_2, ca pereche de laturi opuse ale hexagonului şi punctul Q, ca punct de intersecţie al unei alte perechi de laturi opuse.)

* 24) Două drepte, aflate în întregime în afara foii de hîrtie date, sînt date fiecare de două perechi de drepte care se intersectează în puncte situate în afara foii de hîrtie. Determinaţi punctul lor de intersecţie cu ajutorul unei perechi de drepte care trec prin el.

§ 8. CONICE ŞI SUPRAFEŢE CUADRICE

1. Geometria metrică elementară a conicelor

Pînă acum ne-am ocupat numai de puncte, drepte, plane şi figuri formate din aceste elemente. Dacă geometria proiectivă s-ar reduce la studiul acestor figuri ,,liniare'', ea ar prezenta un interes redus. Este de importanţă fundamentală faptul că geometria proiectivă *nu* se reduce la studiul figurilor liniare, ci include întregul cîmp al secţiunilor conice şi al generalizărilor lor la mai multe dimensiuni. Tratarea metrică de către Apollonius a secţiunilor conice — elipse, hiperbole şi parabole — a fost una dintre marile realizări matematice ale antichităţii. Importanţa secţiunilor conice pentru matematica pură şi aplicată (de exemplu, orbitele planetelor şi ale electronilor din atomul de hidrogen sînt secţiuni conice) nu poate fi preţuită îndeajuns. Nu este de mirare faptul că teoria greacă clasică a secţiunilor conice este încă o parte indispensabilă a învăţămîntului matematic. Însă geometria greacă nu era cîtuşi de puţin încheiată. Două mii de ani mai tîrziu au fost descoperite importantele proprietăţi proiective ale conicelor. În ciuda simplităţii şi frumuseţii acestor proprietăţi, inerţia academică a împiedicat, pînă acum, introducerea lor în programa de liceu.

Vom începe prin a reaminti definiţiile metrice ale secţiunilor conice. Există diferite definiţii de acest fel, a căror echivalenţă este demonstrată în geometria elementară. Cele obişnuite se referă la *focare*. O *elipsă* este definită ca fiind locul geometric al tuturor punctelor P din plan, pentru care suma distanţelor r_1, r_2 la două puncte fixe F_1, F_2, numite focare, are o valoare constantă. (Dacă cele două focare coincid, figura este un cerc.) *Hiperbola* este, prin definiţie, locul geometric al tuturor punctelor P din plan, pentru care valoarea absolută a diferenţei $r_1 - r_2$ este egală cu o constantă. *Parabola* este prin definiţie, locul geometric al tuturor punctelor P, pentru care distanţa r la un punct fix F este egală cu distanţa la o dreaptă dată l.

În limbajul geometriei analitice, aceste curbe pot fi exprimate printr-o ecuație de gradul doi în raport cu coordonatele x, y. Nu este greu de arătat că, reciproc, orice curbă definită analitic printr-o ecuație de gradul doi:

$$ax^2 + by^2 + cxy + dx + ey + f = 0,$$

este fie una din cele trei conice, o dreaptă, o pereche de drepte, un punct, sau este imaginară. Aceasta se face de obicei prin introducerea unui nou sistem de coordonate convenabil, așa cum se procedează în orice curs de geometrie analitică.

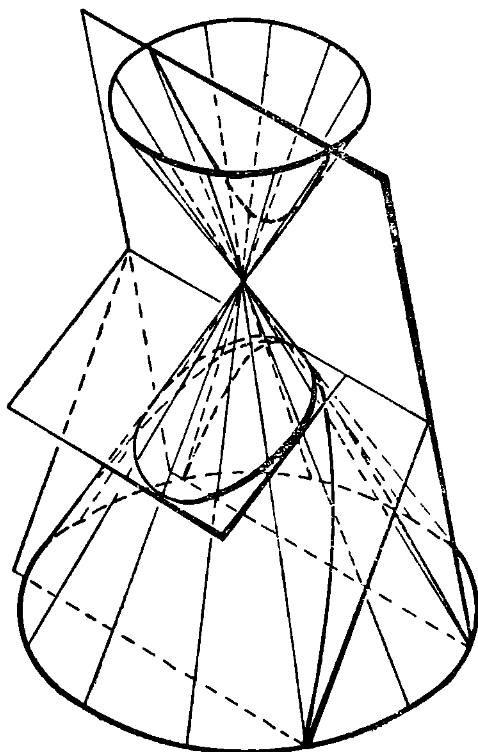


Fig. 94. Secțiuni conice

Aceste definiții ale secțiunilor conice sînt esențial metrice, deoarece ele utilizează noțiunea de distanță. Dar iată încă o definiție care stabilește locul secțiunilor conice în geometria proiectivă: *secțiunile conice sînt proiecțiile unui cerc pe un plan.* Dacă proiectăm un cerc C dintr-un punct O, atunci dreptele

de proiecţie vor forma un con dublu infinit, iar intersecţia acestui con cu un plan π va fi proiecţia lui C. Această intersecţie va fi o elipsă sau o hiperbolă, după cum planul taie una sau ambele pînze ale conului. Cazul intermediar al parabolei apare dacă planul π este paralel cu una dintre dreptele care trec prin O (fig. 94).

Conul de proiecţie nu trebuie să fie neapărat un con circular drept cu vîrful O, aşezat perpendicular deasupra centrului cercului C; el poate fi şi oblic. În toate cazurile, ceea ce vom accepta fără demonstraţie, intersecţia co-

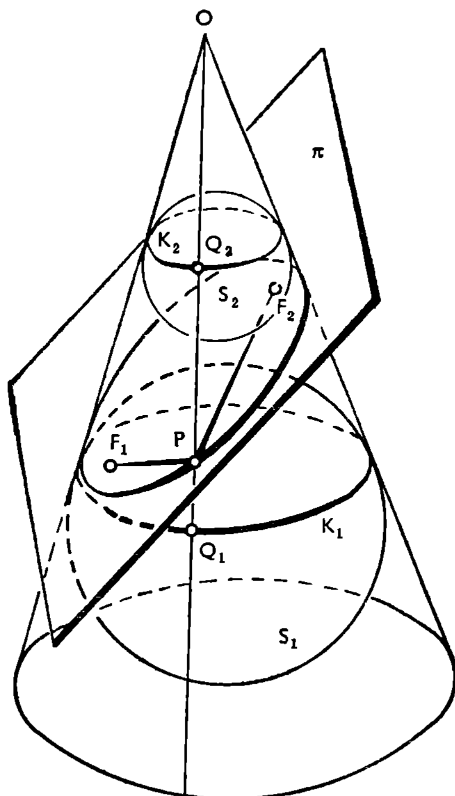


Fig. 95. Sferele lui Dandelin

nului cu un plan va fi o curbă a cărei ecuaţie este de gradul doi, şi reciproc, orice curbă de gradul doi poate fi obţinută dintr-un cerc, printr-o astfel de proiecţie. Acesta este motivul pentru care curbele de gradul doi se numesc secţiuni conice.

Dacă planul intersectează numai o pînză a conului circular drept, atunci, după cum am spus, curba de intersecție E este o elipsă. Putem demonstra că E satisface definiția focală obișnuită a elipsei, dată mai sus, printr-un raționament simplu, dar elegant, dat în 1822 de matematicianul belgian G. P. Dandelin. Demonstrația se bazează pe introducerea a două sfere S_1 și S_2 (fig. 95), care sînt tangente planului π, respectiv în punctele F_1 și F_2, și care ating conul de-a lungul cercurilor paralele K_1 și K_2. Unim un punct arbitrar P al lui E cu F_1 și F_2 și ducem dreapta care unește punctul P cu vîrful O al conului. Această dreaptă se află în întregime pe suprafața conului și intersectează cercurile K_1 și K_2 în punctele Q_1 și Q_2. Acum PF_1 și PQ_1 sînt două tangente din P la S_1, astfel încît

$$PF_1 = PQ_1.$$

În mod analog

$$PF_2 = PQ_2.$$

Adunînd aceste două egalități obținem

$$PF_1 + PF_2 = PQ_1 + PQ_2.$$

Dar $PQ_1 + PQ_2 = Q_1Q_2$ este chiar distanța, pe suprafața conului, între cele două cercuri paralele K_1 și K_2 și, de aceea, este independentă de alegerea particulară a punctului P pe E. Egalitatea obținută,

$$PF_1 + PF_2 = \text{constant}$$

pentru toate punctele P ale lui E, este chiar definiția focală a unei elipse. De aceea, E este o elipsă, iar F_1, F_2 sînt focarele ei.

Exercițiu: Dacă un plan secționează ambele pînze ale conului, atunci curba de intersecție este o hiperbolă. Demonstrați acest lucru, folosind cîte o sferă înscrisă în fiecare pînză a conului.

2. Proprietăți proiective ale conicelor

Pe baza faptelor stabilite în paragraful precedent, vom încerca să adoptăm următoarea definiție: O conică este proiecția unui cerc pe un plan. Această definiție răspunde mai bine spiritului geometriei proiective, în comparație cu definiția focală obișnuită, deoarece ultima se bazează în întregime pe noțiunea metrică de distanță. Chiar ultima definiție dată nu este lipsită de acest defect, deoarece ,,cercul'', este de asemenea o noțiune a geometriei metrice. Vom ajunge îndată la o definiție pur proiectivă a conicelor.

Deoarece am convenit că o conică este pur și simplu proiecția unui cerc (adică, cuvîntul ,,conică'' înseamnă orice curbă din clasa proiectivă a cercului; cf. p. 204), rezultă că orice proprietate a cercului, invariantă prin pro-

iecție, va aparține de asemenea oricărei conice. Se știe că un cerc are proprietatea (metrică) binecunoscută că un arc dat subîntinde același unghi, în orice punct *O* de pe cerc. În fig. 96 unghiul *AOB* subîntins de arcul *AB* este independent de poziția lui *O*. Acest fapt poate fi pus în legătură cu noțiunea proiectivă de biraport, considerînd pe cerc nu două puncte *A*, *B*, ci patru puncte *A*, *B*, *C*, *D*. Cele patru drepte *a*, *b*, *c*, *d* care le unesc cu un al

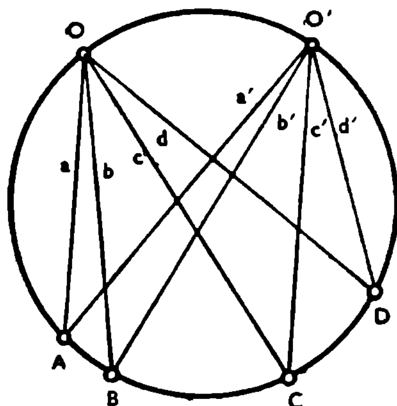


Fig. 96. Biraportul pe un cerc

cincilea punct *O* de pe cerc vor avea un biraport (*abcd*), care depinde numai de unghiurile subîntinse de arcele *CA*, *CB*, *DA*, *DB*. Dacă unim punctele *A*, *B*, *C*, *D* cu un alt punct *O'* de pe cerc, obținem patru drepte *a'*, *b'*, *c'*, *d'*. Din proprietatea cercului, pe care tocmai am menționat-o, cele două cvadruplete de drepte vor fi ,,congruente''[7]. Deci ele vor avea același biraport: $(a'b'c'd') = (abcd)$. Dacă proiectăm acum cercul pe o conică oarecare *K*, vom obține pe *K* patru puncte, pe care le notăm din nou prin *A*, *B*, *C*, *D*, alte două puncte *O*, *O'* și două cvadruplete de drepte *a*, *b*, *c*, *d* și *a'*, *b'*, *c'*, *d'*. Aceste două cvadruplete nu vor fi congruente, deoarece în general egalitatea unghiurilor nu se păstrează prin proiecție. Dar, deoarece biraportul este invariant prin proiecție, egalitatea $(a\ b\ c\ d) = (a'\ b'\ c'\ d')$ va rămîne în vigoare. Acest rezultat ne conduce la o teoremă fundamentală: *Dacă patru puncte oarecare* A, B, C, D *ale unei conice* K *sînt unite cu un al cincilea punct* O *de pe* K, *prin dreptele* a, b, c, d, *atunci valoarea biraportului* (a b c d) *este independentă de poziția lui* O *pe* K (fig. 97).

[7] O mulțime formată din patru drepte concurente *a*, *b*, *c*, *d* se numește congruentă cu o altă mulțime *a'*, *b'*, *c'*, *d'* dacă unghiurile dintre fiecare pereche de drepte din prima mulțime sînt egale, și au același sens ca și unghiurile dintre dreptele corespunzătoare ale celei de-a doua mulțimi.

Acesta este într-adevăr un rezultat remarcabil. Ştim deja că oricare patru puncte de pe o dreaptă dată apar sub acelaşi biraport, oricare ar fi punctul O din care le privim. Această teoremă referitoare la biraport stă la baza geometriei proiective. Acum aflăm că acelaşi rezultat este adevărat pentru patru puncte aflate pe o conică, însă cu o restricţie importantă: al cincilea punct nu mai este arbitrar în plan, dar se poate deplasa pe conica dată.

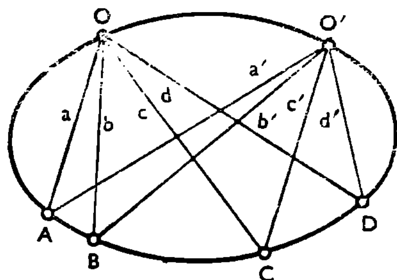


Fig. 97. Birapoarte pe o elipsă

Nu e greu de demonstrat şi teorema reciprocă a acestui rezultat, sub următoarea formă: dacă există două puncte O, O' pe o curbă K astfel încît orice cvadruplet A, B, C, D de puncte de pe K apare sub acelaşi biraport, atît din O cît şi din O', atunci K este o conică (şi de aceea A, B, C, D apar sub acelaşi biraport din orice alt punct O'' al lui K). Omitem demonstraţia.

Aceste proprietăţi proiective ale conicelor sugerează o metodă generală de construire a acestor curbe. Prin *fascicul de drepte* vom înţelege mulţimea tuturor dreptelor din plan, care trec printr-un punct dat O. Să considerăm acum fasciculele determinate de două puncte O şi O', alese pe o conică K. Între dreptele fasciculului O şi acelea ale fasciculului O' putem stabili o corespondenţă biunivocă, asociind o dreaptă a a lui O unei drepte a' a lui O', ori de cîte ori a şi a' se întîlnesc într-un punct A al conicei K. Atunci oricare patru drepte a, b, c, d din fasciculul O vor avea acelaşi biraport, ca şi cele patru drepte corespunzătoare a', b', c', d' din O'. Orice corespondenţă biunivocă dintre două fascicule de drepte, care are această proprietate, se numeşte *corespondenţă proiectivă*. (Evident, această definiţie este duala definiţiei date la p. 197 pentru corespondenţa proiectivă dintre punctele a două drepte.) Două fascicule, între care este definită o corespondenţă proiectivă, se numesc proiectiv legate. Cu această definiţie, putem afirma: Conica K este locul geometric al intersecţiilor dreptelor corespunzătoare din două fascicule proiectiv legate. Această teoremă oferă baza unei definiţii pur proiective a conicelor: *O conică este locul geometric al intersecţiilor dreptelor corespunzătoare din două fascicule proiectiv legate*[8].

[8] În anumite împrejurări, acest loc poate degenera într-o dreaptă (fig. 98).

Este tentant să mergem pe drumul deschis în teoria conicelor de această definiție, însă ne vom limita la cîteva observații.

Se pot obține perechi de fascicule proiectiv legate, în felul următor. Să proiectăm toate punctele P de pe o dreaptă l din două centre diferite, O și O''; să presupunem că în fasciculele de proiecție dreptele a și a'', care intersectează pe l, se corespund. Atunci, cele două fascicule vor fi proiectiv legate. Să luăm

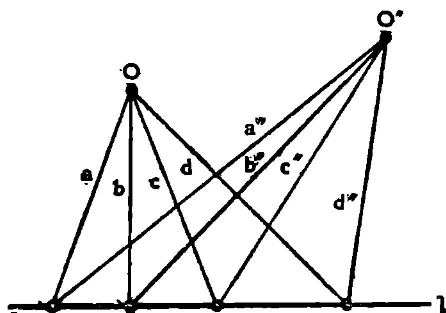


Fig. 98. Preliminariu la construirea fasciculelor proiectiv legate

acum fasciculul O'' și să-l transportăm rigid în altă propoziție O'. Fasciculul rezultat O' va fi proiectiv legat cu O. Mai mult, orice corespondență proiectivă dintre două fascicule poate fi obținută în acest mod. (Acest fapt este dualul exercițiului 1 de la p. 197.) Dacă fasciculele O și O' sînt congruente, obținem un cerc. Dacă unghiurile sîut egale, dar de sens opus, atunci conica este o hiperbolă echilateră (fig. 99).

Observați că accastă definiție a conicei poate da un loc geometric care este o dreaptă, ca în fig. 98. În acest caz, dreapta OO'' se corespunde ei însăși și toate punctele ei contează ca puncte ale locului. Prin urmare, conica degenerează într-o pereche de drepte, ceea ce concordă cu faptul că există secțiuni ale unui con (acelea obținute prin plane duse prin vîrf) care constau din două drepte.

Exerciții: 1) Trasați elipse, hiperbole și parabole cu ajutorul fasciculelor proiective. (Cititorul este îndemnat cu insistență să experimenteze astfel de construcții: Ele vor contribui în mare măsură la înțelegerea lucrurilor.)

2) Sînt date cinci puncte O, O', A, B, C ale unei conice necunoscute K. Se cere să se construiască punctul D, în care o dreaptă dată d, dusă prin O, intersectează pe K. (Indicație: considerați din O dreptele a, b, c, d date de OA, OB, OC și de asemenea, prin O' dreptele a', b', c'. Duceți prin O dreapta d și construiți în O' dreapta d', astfel încît $(abcd) = (a'b'c'd')$. Atunci intersecția dreptelor d și d' este în mod necesar un punct al lui K.)

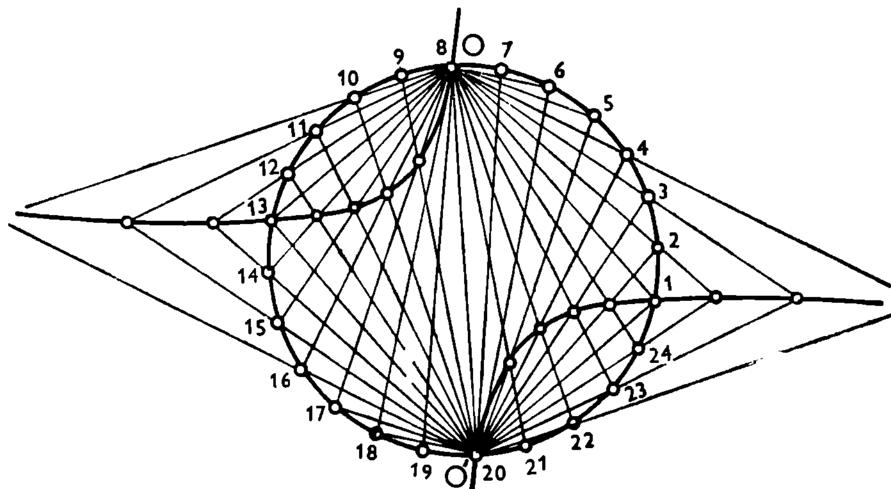


Fig. 99. Cercul și hiperbola echilateră, generate prin fascicule proiective

3. Conicele privite ca curbe tangențiale

Noțiunea de tangentă la o conică aparține geometriei proiective, pentru că tangentă la o conică este o dreaptă care atinge conica într-un singur punct, și această proprietate rămîne neschimbată prin proiecție. Proprietățile proiective ale tangentelor la o conică se bazează pe următoarea teoremă fundamentală: *Biraportul punctelor de intersecție a patru tangente fixe la o conică cu a cincea tangentă este același, oricare ar fi poziția ultimei tangente.*

Demonstrația acestei teoreme este foarte simplă. Deoarece o conică este proiecția unui cerc și deoarece teorema se referă numai la proprietăți invariante prin proiecție, o demonstrație pentru cazul cercului va fi suficientă pentru a stabili teorema în general.

Pentru cerc, teorema aparține geometriei elementare. Fie P, Q, R, S patru puncte pe un cerc K, cu tangentele a, b, c, d; T un alt punct cu tangenta o, intersectată de a, b, c, d în A, B, C, D. Dacă M este centrul cercului, atunci, evident, $\sphericalangle\, TMA = \frac{1}{2} \sphericalangle\, TMP$ și $\frac{1}{2} \sphericalangle\, TMP = \frac{1}{2} \sphericalangle\, \widehat{TP}$. În mod analog, $\sphericalangle\, TMB = \frac{1}{2}\, \widehat{TQ}$ și deci $\sphericalangle\, AMB = \frac{1}{2}\, \widehat{PQ}$, unde $\widehat{PQ}$ este măsura arcului PQ. Deci punctele A, B, C, D sînt proiectate din M cu patru drepte, ale căror unghiuri sînt date de pozițiile fixe ale lui P, Q, R, S. Rezultă că biraportul $(A\ B\ C\ D)$ depinde numai de cele patru tangente a, b, c, d și nu de

poziția particulară a celei de-a cincea tangente o. Aceasta este chiar teorema pe care vroiam s-o demonstrăm.

În paragraful precedent am văzut că o conică poate fi construită prin determinarea punctelor de intersecție ale dreptelor corespunzătoare în două

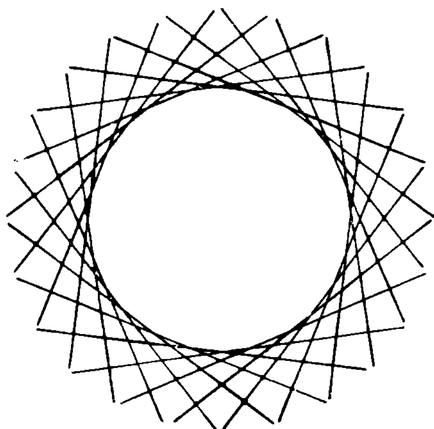


Fig. 100. Cercul ca mulțime de tangente

fascicule proiectiv legate. Teorema pe care tocmai am demonstrat-o ne permite să dualizăm acestă construcție. Să luăm două tangente a și a' ale unei conice K. O a treia tangentă t va intersecta pe a și a', respectiv în două puncte A și A'. Dacă deplasăm tangenta t în jurul conicei, vom stabili în acest mod o corespondență

$$A \leftrightarrow A'$$

între punctele lui a și acelea ale lui a'. Această corespondență dintre punctele lui a și acelea ale lui a' va fi proiectivă, deoarece, în baza teoremei noastre patru puncte oarecare ale lui a vor avea același biraport ca și cele patru puncte corespunzătoare ale lui a'. De aici rezultă că *o conică K*, privită ca mulțime a tangentelor sale, *constă din dreptele care unesc punctele corespunzătoare în două diviziuni punctuale proiectiv legate de pe dreptele* a *și* a'[9].

Acest fapt poate fi folosit pentru a da o definiție proiectivă a conicei, ca

[9] Mulțimea punctelor de pe o dreaptă se numește *diviziune punctuală*. Aceasta este duala fasciculului de drepte.

„curbă de tangente". Să o comparăm cu definiția proiectivă a conicei, dată în paragraful precedent:

I	II
O conică este o mulțime de *puncte* care constă din *punctele de intersecție ale dreptelor corespunzătoare* din două *fascicule* proiectiv *legate*.	**O conică este o mulțime de *drepte*, care constă din *dreptele care unesc punctele corespunzătoare* din două *diviziuni punctuale* proiectiv legate.**

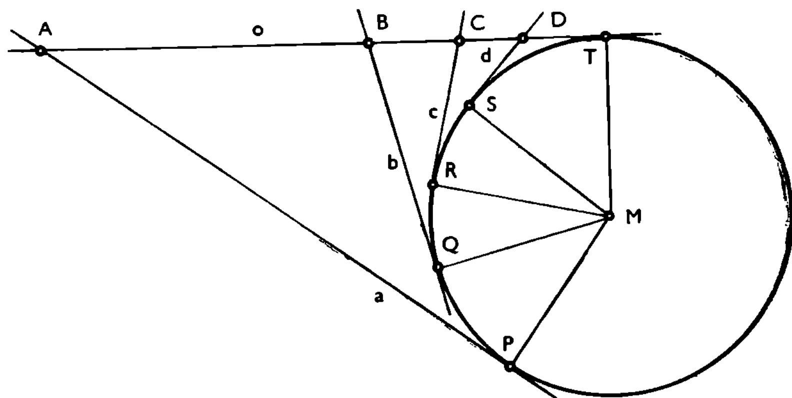


Fig. 101. Proprietatea tangentelor la un cerc

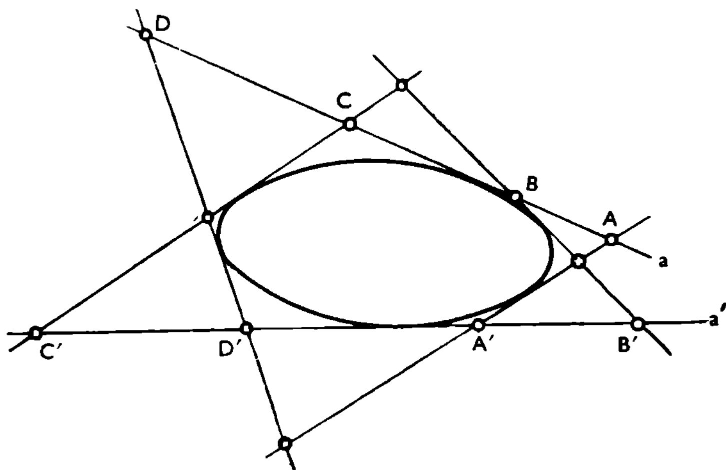


Fig. 102. Diviziuni punctuale proiective pe două tangente la o elipsă

Dacă privim tangenta la o conică într-un punct ca fiind elementul dual al punctului însuși și dacă considerăm o „curbă tangențială" (mulțimea tuturor tangentelor sale) ca fiind duala unei „curbe de puncte" (mulțimea tuturor punctelor ei), atunci deplina dualitate dintre aceste două propoziții este evidentă. În trecerea de la un enunț la altul, înlocuind fiecare noțiune prin duala ei,

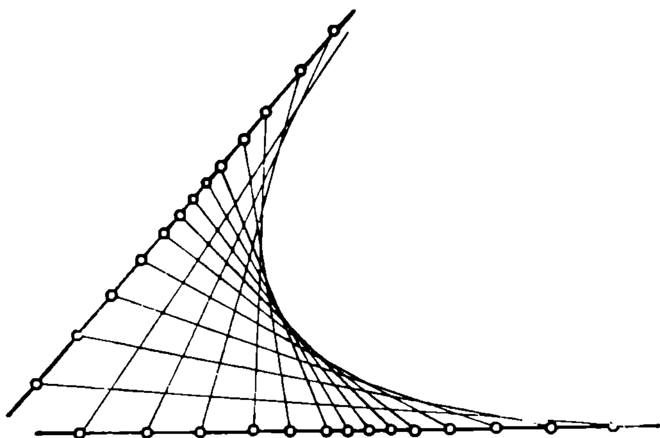


Fig. 103. Parabola definită prin diviziuni congruente

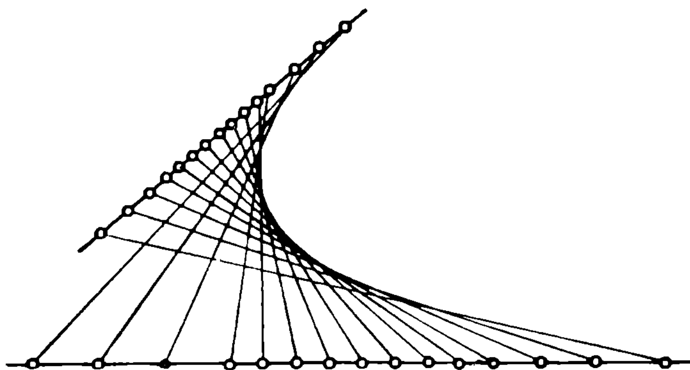


Fig. 104. Parabola definită prin diviziuni asemenea

cuvîntul „conică" rămîne același: într-un caz este o „conică punctuală", definită prin punctele ei; în al doilea, este o „conică tangențială", definită prin tangentele ei (fig. 100).

O consecință importantă a acestui fapt este că principiul dualității din geometria proiectivă plană, enunțat inițial numai pentru puncte și drepte,

poate fi extins acum la conice. *Dacă, în enunțul oricărei teoreme referitoare la puncte, drepte și conice, orice element este înlocuit prin dualul său* (ținînd seama de faptul că dualul unui punct pe o conică este tangenta la conică), *rezultatul va fi tot o teoremă adevărată.* Un exemplu de aplicare a acestui principiu va fi găsit în secțiunea 4 a acestui paragraf.

Construcția conicelor ca curbe tangențiale este arătată în fig. 103—104. Dacă în două diviziuni punctuale proiectiv legate, cele două puncte de la infinit se corespund (așa cum se întîmplă cu diviziunile congruente sau asemenea[10]), atunci conica va fi o parabolă; reciproca este și ea adevărată.

Exercițiu: Demonstrați teorema reciprocă: pe orice două tangente fixate ale unei parabole, o tangentă mobilă determină două diviziuni punctuale asemenea.

4. Teoremele generale ale lui Pascal și Brianchon pentru conice

Una dintre cele mai bune ilustrări ale principiului dualității pentru conice este relația dintre teoremele generale ale lui Pascal și Brianchon. Prima a fost descoperită în 1640, a doua de-abia în 1806. Și totuși fiecare dintre ele este o consecință imediată a celeilalte, deoarece orice teoremă care se referă numai la conice, drepte și puncte rămîne adevărată, dacă este înlocuită prin enunțul dual.

Teoremele enunțate în § 5 sub același nume sînt cazuri degenerate ale următoarelor teoreme mai generale.

Teorema lui Pascal: Laturile opuse ale unui hexagon înscris într-o conică se intersectează în trei puncte coliniare.

Teorema lui Brianchon: Cele trei diagonale care unesc vîrfurile opuse ale unui hexagon circumscris unei conice sînt concurente.

Ambele teoreme au, evident, un caracter proiectiv. Natura lor duală devine mai evidentă, dacă ele sînt formulate în modul următor:

Teorema lui Pascal: Sînt date șase puncte, 1, 2, 3, 4, 5, 6 pe o conică. Să unim punctele succesive prin dreptele (1, 2), (2,3), (3,4), (4,5) , (5,6), (6,1). Să determinăm punctele de intersecție ale lui (1,2) cu (4,5) (2,3) cu (5,6), și (3,4) cu (6,1). Atunci, aceste trei puncte de intersecție se află pe o dreaptă.

Teorema lui Brianchon. Sînt date șase tangente 1, 2, 3, 4, 5, 6 la o conică. Tangentele succesive se intersectează în punctele (1,2), (2,3), (3,4), (4,5), (5,6), (6,1). Duceți dreptele care unesc pe (1,2) cu (4,5), (2,3) cu (5,6) și (3,4) cu (6,1). Atunci aceste drepte trec printr-un punct.

Demonstrațiile pot fi date printr-o particularizare asemănătoare cu aceea folosită în cazurile degenerate. Pentru a demonstra teorema lui Pascal fie

[10] Este evident ce se înțelege prin corespondență „congruentă" sau „asemenea" între două diviziuni punctuale.

A, B, C, D, E, F, vîrfurile unui hexagon înscris într-o conică K. Prin proiecție, putem face ca AB să fie paralelă cu ED, iar FA paralelă cu CD, astfel încît să obținem configurația din fig. 107. (Pentru comoditate, am admis că hexagonul se intersectează cu el însuși, cu toate că acest lucru nu este necesar.) Teorema lui Pascal se reduce acum la afirmația simplă că CB este paralelă cu FE; cu alte cuvinte, dreapta pe care se întîlnesc laturile opuse ale hexago-

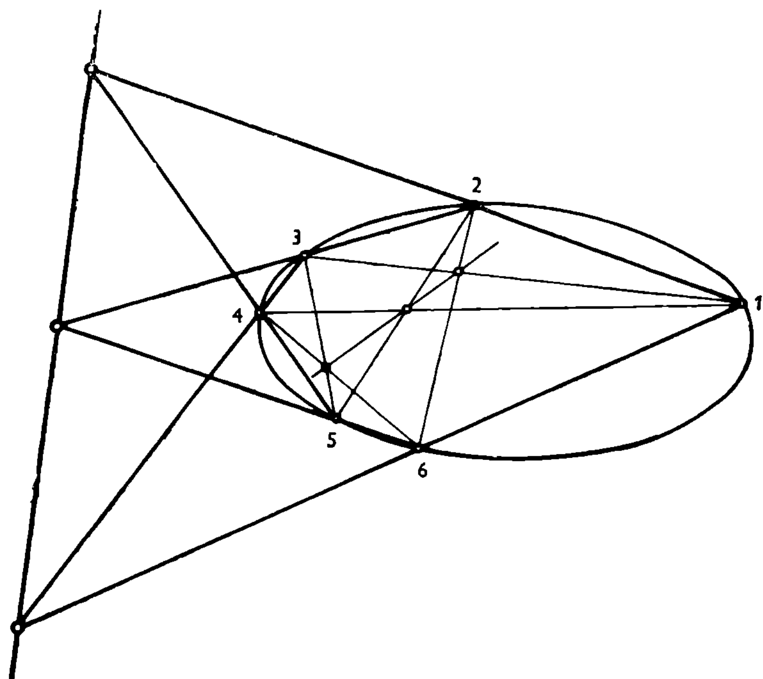


Fig. 105. Configurația generală a lui Pascal. Sînt ilustrate două cazuri: unul pentru hexagonul 1, 2, 3, 4, 5, 6, iar celălalt pentru hexagonul 1, 3, 5, 2, 6, 4

nului este dreapta de la infinit. Pentru a demonstra acest lucru, să considerăm punctele F, A, B, D care, după cum știm, sînt proiectate de drepte care au un biraport constant k, pentru oricare alt punct al lui K; de pildă, pentru C sau E. Să proiectăm aceste puncte din C; atunci dreptele de proiecție intersectează dreapta AF în patru puncte F, A, Y, ∞, care au biraportul k. Deci $YF : YA = k$ (cf. p.203). Dacă aceleași puncte sînt proiectate acum din E pe BA, obținem

$$k = (XAB\infty) = BX : BA.$$

Fig. 106. Configurația generală a lui Brianchon. Din nou sînt ilustrate două cazuri

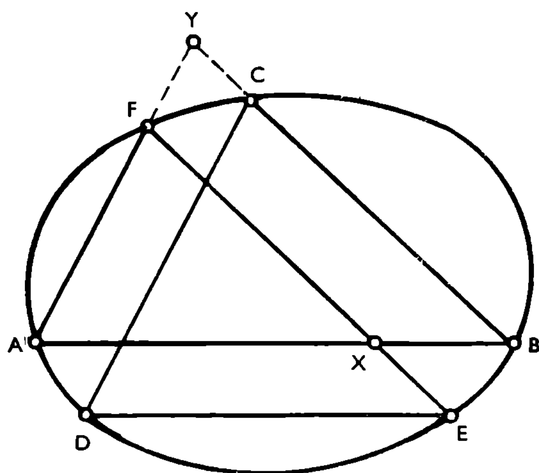


Fig. 107. Demonstrarea teoremei lui Pascal

Deci avem

$$BX : BA = YF : YA,$$

ceea ce stabileşte paralelismul lui YB cu FX. Aceasta completează demonstraţia teoremei lui Pascal.

Teorema lui Brianchon rezultă fie prin aplicarea principiului dualităţii, fie prin raţionament direct, dual aceluia de mai sus. Va fi un bun exerciţiu pentru cititor să efectueze detaliile raţionamentului.

5. Hiperboloidul

În spaţiul cu trei dimensiuni întîlnim aşa-numitele cvadrice (suprafeţe de gradul al doilea); cele mai simple dintre ele sînt sfera şi elipsoidul. Aceste suprafeţe oferă o mai mare varietate şi studiul lor este mai dificil decît al conicelor. Vom discuta pe scurt şi fără a da demonstraţii una dintre cele mai interesante cvadrice, şi anume „hiperboloidul cu o pînză".

Această suprafaţă poate fi definită în felul următor. Alegeţi în spaţiu trei drepte l_1, l_2, l_3, aflate într-o poziţie generală. Prin aceasta înţelegem că două drepte oarecare nu se află în acelaşi plan şi nu sînt toate trei paralele cu un acelaşi plan. Este surprinzător faptul că există o infinitate de drepte în spaţiu, care au proprietatea că fiecare intersectează toate cele trei drepte date. Pentru a vedea aceasta, să ducem un plan oarecare π prin l_1. Atunci π va intersecta pe l_2 şi l_3 în două puncte, iar dreapta m, care uneşte aceste două puncte, va intersecta pe l_1, l_2 şi l_3. Cînd planul π se roteşte în jurul lui l_1, dreapta m se va deplasa, intersectînd mereu pe l_1, l_2, l_3 şi va genera o suprafaţă infinită. Această suprafaţă este hiperboloidul cu o pînză; ea conţine o familie infinită de drepte de tipul m. Trei drepte oarecare din acestea, m_1, m_2, m_3, vor fi, de asemenea, într-o poziţie generală şi toate dreptele din spaţiu care intersectează aceste trei drepte vor fi, de asemenea, pe suprafaţa hiperboloidului. De aici rezultă proprietatea de bază a hiperboloidului: el este format din două familii diferite de drepte; trei drepte oarecare ale aceleiaşi familii sînt într-o poziţie generală şi oricare dreaptă dintr-o familie intersectează toate dreptele celeilalte familii.

O proprietate proiectivă importantă a hiperboloidului este faptul că biraportul celor patru puncte, în care patru drepte date ale unei familii intersectează o dreaptă dată a celeilalte familii, nu depinde de poziţia ultimei drepte. Acest lucru rezultă direct din metoda de construire a hiperboloidului cu ajutorul planului care se roteşte, după cum cititorul ar putea să arate ca exerciţiu.

Una dintre proprietăţile remarcabile ale hiperboloidului este faptul că deşi el conţine două familii de drepte, care se intersectează, aceste drepte nu fac suprafaţa rigidă. Dacă confecţionăm un model al suprafeţei din fire de sîrmă, articulate la intersecţii, atunci întreaga figură poate fi deformată în mod continuu, în diferite moduri.

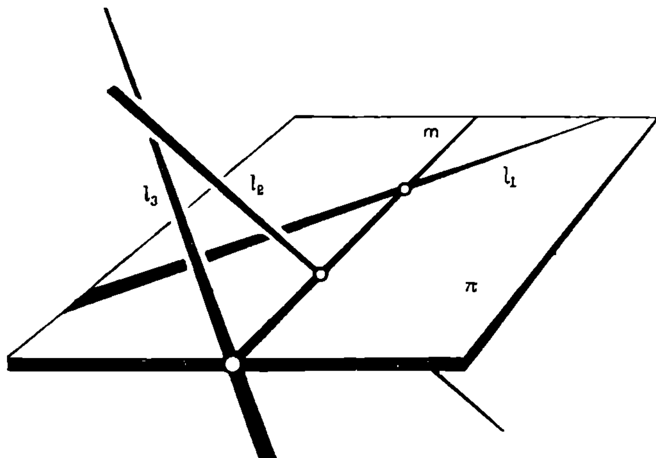


Fig. 108. Construirea dreptelor care intersectează trei drepte fixe, aflate în poziție generală

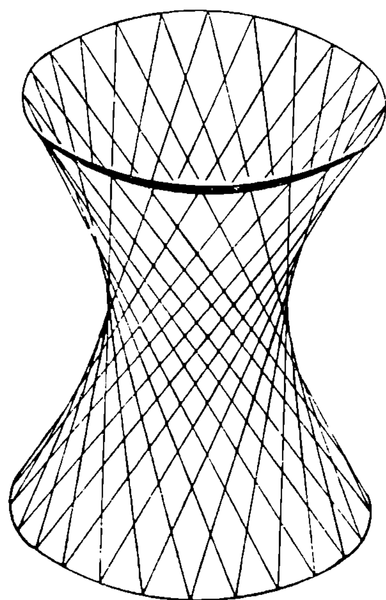


Fig. 109. Hiperboloidul

§ 9. AXIOMATICA ȘI GEOMETRIA NEEUCLIDIANĂ

1. Metoda axiomatică

Metoda axiomatică în matematică începe cel puțin cu Euclid. Ar fi însă complet greșit să presupunem că matematica greacă a fost dezvoltată sau prezentată exclusiv în forma axiomatică, rigidă, a *Elementelor*. Dar impresia produsă de această operă asupra generațiilor următoare a fost atît de mare, încît ea a devenit un model pentru orice demonstrație riguroasă în matematică. Uneori chiar filozofi, ca de pildă Spinoza, în *Ethica, more geometrico demonstrata*, a încercat să prezinte raționamentele sub forma unor teoreme deduse din definiții și axiome. În matematica modernă, după o abatere de la tradiția euclidiană, care a avut loc în decursul secolelor XVII și XVIII, a urmat o pătrundere din ce în ce mai mare a metodei axiomatice în toate domeniile. Unul dintre cele mai recente rezultate a fost crearea unei noi discipline, logica matematică.

În termeni generali, punctul de vedere axiomatic poate fi descris în modul următor : A demonstra o teoremă într-un sistem deductiv înseamnă a arăta că teorema este o consecință logică, necesară, a unor propoziții demonstrate anterior ; acestea, la rîndul lor, trebuie să fie demonstrate, și așa mai departe. De aceea, procesul demonstrării matematice s-ar reduce la sarcina imposibilă a unei succesiuni infinite de demonstrații, cu excepția cazului în care, întorcîndu-ne în urmă, avem voie să ne oprim într-un anumit punct. Prin urmare, trebuie să existe un număr de propoziții, numite *postulate* sau *axiome*, care sînt acceptate ca adevărate și pentru care nu se cer demonstrații. Din acestea putem încerca să deducem toate celelalte teoreme prin raționament pur logic. Dacă faptele unui domenin științific sînt ordonate logic în acest mod, încît totul poate fi dedus dintr-un număr de propoziții alese (de preferință în număr redus, simple și plauzibile), atunci se spune că domeniul este prezentat sub o formă axiomatică. Alegerea propozițiilor care joacă rolul de axiome este în mare măsură arbitrară. Însă nu cîștigăm prea mult prin metoda axiomatică, dacă postulatele nu sînt simple și sînt prea numeroase. Mai mult, postulatele trebuie să fie *consistente*, în sensul că două teoreme oarecare, deductibile din ele, nu trebuie să fie contradictorii, și *complete*, astfel încît orice teoremă a sistemului să poată fi dedusă din axiome. Din motive de economie, mai este de dorit, ca postulatele să fie *independente*, în sensul că nici unul dintre ele să nu fie o consecință logică a celorlalte. Problema consistenței și a completitudinii unui sistem de axiome a fost mult controversată. Diferite convingeri filozofice referitoare la ultimele rădăcini ale cunoașterii umane au dus la puncte de vedere ireconciliabile asupra fundamentelor matematicii. Dacă entitățile matematice sînt considerate ca obiecte substanțiale din domeniul „intuiției pure”, independente de definițiile și actele individuale ale intelectului uman, atunci desigur, nu pot exista contradicții, deoarece faptele matematice sînt propoziții obiectiv adevărate, care descriu realități existente. Din acest punct de vedere,

„kantian", nu se pune problema consistenței. Din nefericire însă, corpul real al matematicii nu poate fi potrivit într-un cadru filozofic atît de simplu. Intuiționiștii matematici moderni nu se bazează pe intuiția pură, în sensul larg al lui Kant. Ei acceptă infinitul numărabil drept copil legitim al intuiției și admit doar proprietăți constructive; dar în acest fel, noțiuni fundamentale, ca de pildă continuul numeric, ar fi eliminate, părți importante ale matematicii actuale ar fi excluse, iar restul s-ar complica fără speranță.

Cu totul deosebit este punctul de vedere adoptat de „formaliști". Ei nu atribuie o realitate intuitivă obiectelor matematice și nici nu pretind că axiomele exprimă adevăruri evidente, referitoare la realitățile intuiției pure; ei se ocupă numai de procedeul logic formal al raționamentului, pe baza postulatelor. Această atitudine prezintă un avantaj cert față de intuiționism, deoarece ea lasă matematicii toată libertatea necesară pentru teorie și aplicații. Dar ea impune formalistului obligația de a demonstra că axiomele lui, care apar acum ca creații arbitrare ale minții omenești, nu pot duce la o contradicție. Au fost făcute mari eforturi în cursul ultimilor 20 de ani, pentru a găsi aceste demonstrații de consistență, cel puțin pentru axiomele aritmeticii și ale algebrei, și pentru conceptul de continuu numeric. Rezultatele sînt deosebit de însemnate, dar un succes complet este încă îndepărtat. Într-adevăr, rezultatele recente indică faptul că astfel de eforturi nu pot avea un succes deplin, în sensul că demonstrațiile pentru consistență și completitudine nu sînt posibile, rămînînd strict în interiorul unor sisteme închise de concepte. Este destul de remarcabil faptul că toate aceste raționamente referitoare la fundamente utilizează metode care, în sine, sînt constructive și sînt ghidate de modele intuitive.

Accentuată de paradoxurile teoriei mulțimilor (cf. p. 103), ciocnirea dintre intuiționiști și formaliști a devenit de notorietate publică datorită unor adepți pasionați ai acestor școli. Lumea matematică a răsunat de strigătul „criză în fundamente". Dar alarma nu a fost și nici nu trebuie să fie luată în serios. Acordînd toată încrederea rezultatelor obținute în lupta pentru clarificarea fundamentelor, ar fi cu totul nedrept să tragem concluzia că corpul viu al matematicii ar fi amenințat de astfel de divergențe de păreri sau de paradoxurile inerente unei tendințe necontrolate spre generalitatea fără margini.

Deosebit de considerațiile filozofice și de interesul pentru fundamente, abordarea axiomatică a unui subiect matematic este calea naturală de descoperire a rețelei de legături dintre diferitele fapte și de a indica scheletul logic esențial al structurii. Se întîmplă uneori ca o astfel de concentrare asupra structurii formale, renunțînd la înțelesul intuitiv al noțiunilor, să facă mai ușoară găsirea generalizărilor și a aplicațiilor, care ar putea fi pierdute din vedere la o tratare mai intuitivă. Însă o descoperire însemnată sau o privire clarificatoare se obține rareori numai prin procedeul axiomatic. Gîndirea constructivă, ghidată de intuiție, este adevărata sursă a dinamicii matematicii. Cu toate că forma axiomatică este un ideal, este o iluzie periculoasă să ne închipuim că axiomatica constituie *esența* matematicii. Intuiția constructivă a matematicianului con-

feră matematicii elemente nedeductive și iraționale, care o fac comparabilă cu muzica și arta.

Din vremea lui Euclid, geometria a fost prototipul oricărei discipline axiomatizate. Timp de secole, sistemul de axiome al lui Euclid a fost obiectul unui studiu intens. Dar numai de curînd a devenit evident faptul că aceste postulate trebuie să fie modificate și completate, dacă vrem ca toată geometria elementară să poată fi dedusă din ele. De-abia în secolul al XIX-lea, de pildă, Pasch a descoperit că ordonarea punctelor pe o dreaptă, noțiunea ,,între", are nevoie de o axiomă specială. Pasch a formulat propoziția următoare sub forma unei axiome: o dreaptă care intersectează o latură a unui triunghi într-un punct, altul decît vîrfurile, mai intersectează încă o latură. (Neglijarea acestor amănunte duce la multe paradoxuri evidente, în care consecințe absurde — ca de pildă binecunoscuta ,,demonstrație" că orice triunghi este isoscel — par a fi deduse riguros din axiomele lui Euclid. Aceasta se face, de obicei, cu ajutorul unei figuri desenate greșit, în care anumite drepte par a se intersecta în interiorul sau în exteriorul anumitor triunghiuri sau cercuri, ceea ce nu se întîmplă în realitate.)

În celebra sa carte, *Grundlagen der Geometrie* (prima ediție a fost publicată în 1901), Hilbert a dat un sistem satisfăcător de axiome pentru geometrie și, în același timp, a făcut un studiu exhaustiv al independenței, consistenței și completitudinii lui.

În orice sistem de axiome trebuie să intervină anumite concepte nedefinite, ca de pildă ,,punctul" și ,,dreapta" din geometrie. ,,Înțelesul" sau legătura lor cu obiectele lumii fizice nu este esențial pentru *matematică*. Ele pot fi privite ca entități pur abstracte, ale căror proprietăți matematice în sistemul deductiv sînt date, în întregime, prin relațiile stabilite între ele în enunțul axiomelor. De exemplu, în geometria proiectivă am putea începe cu conceptele nedefinite de ,,punct", ,,dreaptă" și ,,incidență" și cu două axiome duale: ,,oricare două puncte distincte sînt incidente cu o singură dreaptă" și ,,oricare două drepte distincte sînt incidente cu un singur punct". Din punct de vedere axiomatic, tocmai forma duală a acestor axiome este izvorul principiului dualității din geometria proiectivă. Orice teoremă care conține în enunțul și în demonstrația ei numai elemente legate prin axiome duale trebuie să admită o dualizare, deoarece demonstrația teoremei inițiale constă în aplicarea succesivă a anumitor axiome, și aplicarea axiomelor duale în aceeași ordine va produce o demonstrație a teoremei duale.

Mulțimea axiomelor geometriei constituie *definiția implicită* a tuturor noțiunilor geometrice ,,nedefinite": ,,punct", ,,dreaptă", ,,incident" etc. Pentru aplicații este important faptul că noțiunile și axiomele geometriei să corespundă bine propozițiilor verificabile din punct de vedere fizic, referitoare la obiecte palpabile, ,,reale". Realitatea fizică, aflată în spatele noțiunii de ,,punct", este aceea a unui obiect foarte mic, ca de pildă semnul făcut cu creionul, în timp ce ,,dreapta" este o abstracție a firului întins sau a razei de lu-

mină. Proprietățile acestor puncte și drepte fizice concordă experimental, mai mult sau mai puțin, cu axiomele formale ale geometriei. Este de conceput ca experimente mai precise să necesite modificarea acestor axiome, dacă ele trebuie să descrie în mod adecvat fenomenele fizice. Dar dacă axiomele formale nu ar concorda mai mult sau mai puțin cu proprietățile obiectelor fizice, atunci geometria ar prezenta un interes limitat. Astfel, chiar pentru formaliști, există ceva care are o influență mai mare asupra matematicii decît rațiunea umană.

2. Geometria neeuclidiană hiperbolică

Există o axiomă a geometriei euclidiene, al cărei „adevăr", adică a cărei corespondență cu datele empirice referitoare la firele întinse sau la razele de lumină nu mai este evidentă. Aceasta este celebrul *postulat al paralelei unice*, care afirmă că prin orice punct exterior unei drepte se poate duce o *singură* dreaptă paralelă cu dreapta dată. Caracterul remarcabil al acestei axiome constă în faptul că ea enunță o proprietate a dreptei în *întregime*, imaginată ca fiind nelimitată în ambele sensuri; într-adevăr, a spune că două drepte sînt paralele înseamnă că ele nu se intersectează, oricît le-am prelungi. Este evident faptul că există multe drepte care trec printr-un punct și care nu intersectează o dreaptă dată, într-un *cadru finit fixat*, oricît de mare. Deoarece lungimea maximă posibilă a unei rigle, a unui fir sau chiar a unei raze de lumină vizibile cu telescopul este desigur finită și deoarece în interiorul oricărui cerc finit există o infinitate de drepte care trec printr-un punct dat și nu intersectează o dreaptă dată în interiorul cercului, rezultă că această axiomă nu poate fi verificată niciodată prin experiență. Toate celelalte axiome ale geometriei euclidiene au un caracter finit, prin faptul că ele se referă la porțiuni finite ale dreptelor sau la figuri plane de întindere finită. Faptul că axioma paralelelor nu este verificabilă din punct de vedere experimental ridică problema dacă ea este, sau nu, *independentă* de celelalte axiome. Dacă ea ar fi o consecință logică necesară a celorlalte, atunci am putea să o scoatem din rîndul axiomelor și să o demonstrăm cu ajutorul celorlalte axiome euclidiene. Timp de secole, matematicienii au încercat să găsească o astfel de demonstrație, datorită părerii larg răspîndite printre cei care studiau geometria, că postulatul paralelelor are un caracter esențial diferit de cel al celorlalte postulate, lipsindu-i plauzibilitatea convingătoare pe care ar trebui să o aibă o axiomă a geometriei. Una dintre primele încercări de această natură a fost făcută de Proclus (secolulul IV-lea e.n), un comentator al lui Euclid, care a încercat să evite folosirea unui postulat special al paralelelor, *definind* paralela la o dreaptă dată ca fiind locul geometric al tuturor punctelor aflate la o distanță dată de dreapta dată. Prin aceasta, Proclus a scăpat din vedere faptul că, în acest mod, dificultatea a fost deplasată în alt loc, deoarece în acest caz ar fi fost necesar să se demonstreze că locul geometric al acestor puncte este o dreaptă. Deoarece Proclus nu a putut demonstra acest lucru, el a trebuit să-l accepte ca postulat, în locul axiomei paralelelor și

în acest mod nu s-a cîștigat nimic, deoarece se vede cu ușurință că cele două axiome sînt echivalente. Iezuitul Saccheri (1667—1733) și mai tîrziu Lambert (1728—1777) au încercat să demonstreze postulatul paralelelor prin metoda indirectă, presupunînd contrariul și căutînd să obțină consecințe absurde. Departe de a fi absurde, concluziile lor erau de fapt teoreme de geometrie neeuclidiană, dezvoltată mai tîrziu. Dacă nu le-ar fi privit ca absurdități, ci ca propoziții de sine stătătoare, ei ar fi fost descoperitorii geometriei neeuclidiene.

Pe atunci, orice sistem geometric, care nu era în acord deplin cu cel al lui Euclid, ar fi fost considerat un nonsens evident. Kant, cel mai influent filozof al vremii, a formulat această atitudine, afirmînd că axiomele lui Euclid sînt inerente intelectului uman, și de aceea au o valabilitate obiectivă pentru spațiul „real". Această încredere în axiomele geometriei euclidiene ca exprimînd adevăruri incontestabile, existente în domeniul intuiției pure, a fost unul dintre principiile fundamentale ale filozofiei lui Kant. Dar cu timpul, nici vechile moduri de gîndire, nici autoritatea filozofică nu au putut suprima convingerea că șirul nesfîrșit de eșecuri suferite în căutarea unei demonstrații pentru postulatul paralelelor nu se datora unei lipse de ingeniozitate, ci mai degrabă faptului că postulatul paralelelor este *independent* de celelalte. (În mod asemănător, lipsa succesului în demonstrarea faptului că ecuația generală de gradul V ar putea fi rezolvată prin radicali a dus la bănuiala, verificată mai tîrziu, că o astfel de rezolvare este imposibilă.) Matematicianul maghiar Bolyai (1802—1860) și matematicianul rus Lobacevski (1793—1856) au rezolvat problema, construind în toate amănuntele o geometrie în care axioma paralelelor nu este valabilă. Atunci cînd entuziastul și genialul tînăr Bolyai a prezentat lucrarea sa lui Gauss, „prințul matematicienilor", pentru recunoașterea pe care o aștepta cu atîta nerăbdare, el a fost informat că lucrarea lui a fost anticipată de Gauss însuși, dar că acesta nu s-a îngrijit de publicarea rezultatelor obținute, fiindu-i teamă de o publicitate prea zgomotoasă.

Ce semnificație are independența postulatului paralelelor? Aceasta înseamnă că este posibil să se construiască un sistem consistent de propoziții „geometrice", referitoare la puncte, drepte etc., deduse dintr-un sistem de axiome în care postulatul paralelelor este înlocuit cu negația lui. Un astfel de sistem se numește geometrie neeuclidiană. A fost necesar curajul intelectual al lui Gauss, Bolyai și Lobacevski, pentru a recunoaște că o astfel de geometrie, bazată pe un sistem neeuclidian de axiome, poate fi consistentă.

Pentru a demonstra consistența noii geometrii, nu este suficient să deducem un mare număr de teoreme neeuclidiene, așa cum au făcut Bolyai și Lobacevski. În schimb, am învățat să construim „modele" ale unei astfel de geometrii, care satisfac toate axiomele geometriei lui Euclid, cu excepția postulatului paralelelor. Cel mai simplu model de acest fel a fost dat de Felix Klein, a cărui operă în acest domeniu a fost stimulată de ideile geometrului englez Cayley (1821—1895). În acest model pot fi trasate o infinitate de „drepte", „paralele" unei

drepte date, care trec printr-un punct exterior acesteia din urmă. O astfel de geometrie se numeşte geometrie Bolyai-Lobacevski sau geometrie „hiperbolică". (Motivul acestei denumiri va fi dat la p. 243.)

Modelul lui Klein se construieşte considerînd mai întîi obiecte ale geometriei euclidiene obişnuite, *renumind* apoi unele obiecte şi relaţiile dintre ele, în aşa fel încît să rezulte o geometrie neeuclidiană. Aceasta trebuie să fie tot atît de consistentă ca şi geometria euclidiană iniţială, deoarece ea ne este prezentată din alt punct de vedere şi este descrisă cu alte cuvinte decît un sistem de fapte ale geometriei euclidiene obişnuite. Acest model poate fi înţeles cu uşurinţă, cu ajutorul cîtorva noţiuni de geometrie proiectivă.

Dacă supunem planul unei transformări proiective pe un alt plan sau mai degrabă pe el însuşi (făcînd apoi ca planul-imagine să coincidă cu planul iniţial), atunci, în general, un cerc şi interiorul său vor fi transformate într-o secţiune conică. Însă se poate arăta cu uşurinţă (aici omitem demonstraţia) că există o infinitate de transformări proiective ale planului pe el însuşi, astfel încît un cerc împreună cu interiorul său să fie transformate în ele însele. Prin astfel de transformări, punctele interiorului sau ale frontierei sînt deplasate în general în alte poziţii, însă rămîn în interiorul sau pe frontiera cercului. (De fapt, putem deplasa centrul cercului în orice alt punct interior.) Să considerăm mulţimea tuturor transformărilor de acest fel. Desigur, ele nu vor lăsa invariante formele figurilor, şi de aceea nu sînt deplasări rigide în sensul obişnuit al cuvîntului. Însă acum facem pasul hotărîtor şi le *numim* „deplasări neeuclidiene" ale geometriei pe care o construim. Cu ajutorul acestor „deplasări" putem defini congruenţa — două figuri fiind *numite* congruente, dacă există o deplasare neeuclidiană care o transformă pe una în cealaltă.

Modelul lui Klein al geometriei hiperbolice este următorul: „planul" constă numai din punctele interioare cercului; punctele exterioare sînt neglijate. Orice punct interior cercului se *numeşte* „punct" neeuclidian; orice coardă a cercului se *numeşte* „dreaptă" neeuclidiană; „deplasarea" şi „congruenţa" sînt definite ca mai sus; unirea „punctelor" şi găsirea intersecţiei „dreptelor" în sensul neeuclidian înseamnă acelaşi lucru ca şi în geometria euclidiană. Este uşor de arătat că noul sistem satisface toate postulatele geometriei euclidiene, cu singura excepţie a postulatului paralelelor. Faptul că postulatul paralelelor nu este în vigoare în noul sistem se arată observînd că prin orice „punct" care nu se află pe o „dreaptă" se pot duce o infinitate de „drepte", care nu au nici un „punct" comun cu „dreapta" dată. Prima „dreaptă" este o coardă euclidiană a cercului, în timp ce a doua „dreaptă" poate fi oricare din corzile care trec prin „punctul" dat şi nu intersectează prima „dreaptă" în interiorul cercului. Acest model simplu este pe deplin suficient pentru rezolvarea problemei fundamentale care a dat naştere geometriei neeuclidiene: el demonstrează că postulatul paralelelor nu poate fi dedus din celelalte axiome ale geometriei euclidiene. Presupunînd că ar putea fi dedus din celelalte axiome, el ar fi o teoremă adevărată în geometria modelului lui Klein, ceea ce nu este adevărat.

La drept vorbind, acest raționament se bazează pe ipoteza că geometria modelului lui Klein este consistentă, astfel încît nu se poate demonstra o teoremă ca și contrara ei.

Dar geometria modelului lui Klein este desigur tot atît de consistentă ca și geometria euclidiană obișnuită, deoarece propozițiile referitoare la „puncte", „drepte" etc. din modelul

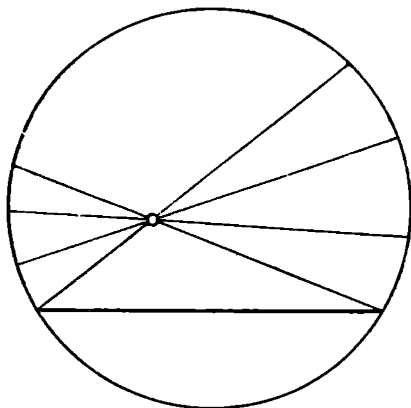


Fig. 110. Modelul neeuclidian al lui Klein

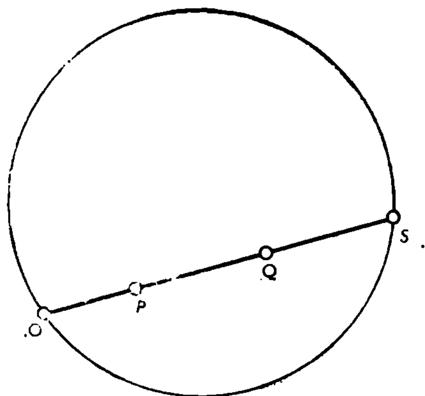


Fig. 111. Distanța neeuclidiană

lui Klein sînt doar diferite moduri de exprimare a anumitor teoreme ale geometriei euclidiene. O demonstrație satisfăcătoare a consistenței axiomelor geometriei euclidiene nu a fost dată niciodată, cu excepția reducerii acestei probleme la noțiunile geometriei analitice, și deci la acelea ale continuului numeric, a cărui consistență a rămas o problemă deschisă.

Ar trebui să menționăm aici încă un amănunt care depășește scopul nostru imediat, și anume : cum se poate defini „distanța" neeuclidiană în modelul lui Klein. Această „distanță" trebuie să fie invariantă față de orice „deplasare" neeuclidiană, pentru că deplasarea ar trebui să lase distanțele invariante. Știm că birapoartele sînt invariante prin proiecție. Un biraport, care se referă la două puncte arbitrare P și Q din interiorul cercului, apare imediat dacă prelungim segmentul PQ pînă la întîlnirea cu cercul, în punctele O și S. Biraportul $(OSQP)$ ale acestor patru puncte este un număr (pozitiv) pe care am putea spera să îl luăm ca definiție a „distanței" PQ, dintre P și Q. Însă această definiție trebuie modificată puțin, pentru a o face operabilă, și aceasta deoarece dacă trei puncte P, Q, R se află pe o dreaptă, ar trebui să avem $\overline{PQ} + \overline{QR} = \overline{PR}$. Dar, în general,

$$(OSQP) + (OSRQ) \neq (OSRP).$$

În schimb, avem relația

$$(1) \quad (OSQP)\cdot(OSRQ) = (OSRP),$$

după cum rezultă din egalitățile

$$(OSQP)(OSRQ) = \left\{\frac{QO/QS}{PO/PS}\right\}\cdot\left\{\frac{RO/RS}{QO/QS}\right\} = \frac{RO/RS}{PO/PS} = (OSRP).$$

Ca o consecință a egalității (1), putem da o definiție aditivă satisfăcătoare, măsurînd „distanța" nu prin biraport, ci prin *logaritmul biraportului*:

$\overline{PQ}$ = distanța neeuclidiană de la P la $Q = \log\,(OSQP)$.

Distanța va fi un număr pozitiv, deoarece $(OSQP) > 1$ dacă $P \neq Q$. Folosind proprietatea fundamentală a logaritmului (cf. p. 462), din (1) rezultă că $\overline{PQ} + \overline{QR} = \overline{PR}$. Baza aleasă pentru logaritm nu are importanță, deoarece schimbarea bazei schimbă doar unitatea de măsură. În treacăt fie zis, dacă unul dintre puncte, de pildă Q, se apropie de cerc, atunci distanța neeuclidiană $\overline{PQ}$ tinde spre infinit. Acest lucru arată că o dreaptă a geometriei noastre neeuclidiene este de lungime neeuclidiană infinită, cu toate că în sensul euclidian obișnuit ea este doar un segment finit al unei drepte.

3. Geometria și realitatea

Modelul lui Klein arată că geometria hiperbolică, privită ca sistem deductiv formal, este tot atît de consistentă ca și geometria euclidiană clasică. Se pune atunci problema, care din două trebuie preferată ca descriere a geometriei lumii fizice? După cum am mai văzut, experiența nu poate decide niciodată dacă există o singură dreaptă sau o infinitate de drepte, care trec printr-un punct și sînt paralele cu o dreaptă dată. În geometria euclidiană însă, suma unghiurilor oricărui triunghi este egală cu 180°, în timp ce în geometria hiperbolică, suma este întotdeauna mai mică decît 180°. În consecință, Gauss a efectuat o experiență pentru a rezolva problema. El a măsurat cu precizie unghiurile unui triunghi, format de vîrfurile a trei munți destul de îndepărtați unul de altul, și a găsit că suma unghiurilor este egală cu 180°, în limitele erorii experimentale. Dacă rezultatul ar fi fost mai mic decît 180°, consecința ar fi fost că geometria hiperbolică este preferabilă pentru descrierea realității fizice. Dar, după cum s-a arătat prin această experiență, problema nu a putut fi rezolvată, deoarece pentru triunghiurile mici, ale căror laturi au lungimea de numai cîteva mile, abaterea față de 180° în geometria hiperbolică ar putea fi atît de mică, încît să nu fi fost detectabilă cu instrumentelelui Gauss. Astfel, deși experimentul nu a dus la nici o concluzie, el a arătat că geometriile euclidiană și hiperbolică, care se deosebesc mult din punct de vedere *global*, coincid atît de mult pentru figuri relativ mici, încît, din punct de vedere experimental, ele sînt echivalente. De aceea, atîta timp cît se iau în considerare proprie-

tățile pur *locale* ale spațiului, alegerea dintre cele două geometrii trebuie făcută numai pe baza simplității și comodității. Deoarece sistemul euclidian este mai ușor de mînuit, îl vom folosi atîta timp cît considerăm distanțe mici (de cîteva milioane de mile!). Nu trebuie totuși să ne așteptăm ca el să fie în mod necesar potrivit pentru descrierea universului în ansamblu, sub aspectele sale globale. Aici, situația este cu totul analogă cu aceea care există în fizică, unde sistemele lui Newton și Einstein dau aceleași rezultate pentru distanțe și viteze mici, dar diferă atunci cînd intervin mărimi foarte mari.

Importanța revoluționară a descoperirii geometriei neeuclidiene constă în faptul că ea a spulberat ideea că axiomele lui Euclid, ar fi un sistem matematic imuabil, în care trebuie să fie încadrată cunoașterea experimentală a realității fizice.

4. Modelul lui Poincaré

Matematicianul este liber să considere o „geometrie" ca fiind definită de orice sistem de axiome consistente, referitoare la „puncte", „drepte" etc. Investigațiile sale vor fi folositoare fizicianului, numai dacă aceste axiome corespund comportării fizice a obiectelor din lumea reală. Din acest punct de vedere, dorim să examinăm înțelesul propoziției „lumina se propagă în linie dreaptă". Dacă aceasta este privită ca *definiție fizică* a „dreptei", atunci axiomele geometriei trebuie să fie alese în așa fel, încît să corespundă comportării razelor de lumină. Să ne imaginăm, o dată cu Poincaré, o lume formată din interiorul unui cerc *C*, astfel încît viteza luminii în orice punct din interiorul cercului să fie proporțională cu distanța punctului la circumferință. Se poate arăta că razele de lumină vor avea atunci forma unor arce de cerc, ortogonale circumferinței *C*. Într-o astfel de lume, proprietățile geometrice ale „dreptelor" (definite ca raze de lumină), se vor deosebi de proprietățile euclidiene ale dreptelor. În particular, axioma paralelelor nu va rămîne în vigoare, pentru că vor exista o infinitate de drepte care trec printr-un punct și care nu intersectează o „dreaptă" dată. De fapt, „punctele" și „dreptele" acestei lumi vor avea proprietățile geometrice ale „punctelor" și „dreptelor" modelului lui Klein. Cu alte cuvinte, vom avea un alt model al geometriei hiperbolice. Însă geometria lui Euclid va fi și ea aplicabilă acestei lumi; în loc de a fi „drepte" neeuclidiene, razele de lumină vor fi cercuri euclidiene, ortogonale lui *C*. Astfel, vedem că diferite sisteme ale geometriei pot descrie aceeași situație fizică, cu condiția ca obiectele fizice (în cazul nostru, razele de lumină) să fie corelate cu diferitele concepte ale celor două sisteme:

rază de lumină → „dreaptă" — geometria hiperbolică

rază de lumină → „cerc" — geometria euclidiană.

Deoarece noțiunea de dreaptă a geometriei euclidiene corespunde comportării unei raze de lumină într-un mediu omogen, am spune că geometria regiunii din

interiorul lui C este hiperbolică, înțelegînd doar că proprietățile fizice ale razelor de lumină corespund în această lume proprietăților „dreptelor” geometriei hiperbolice.

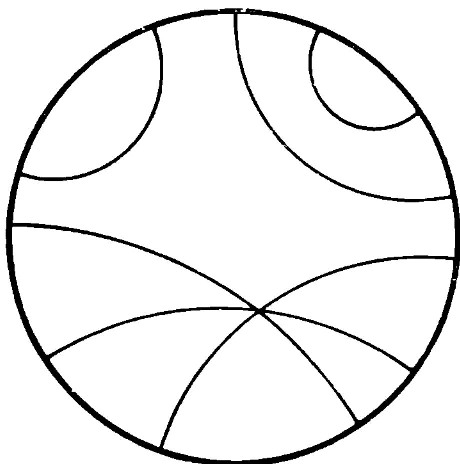


Fig. 112. Modelul neeuclidian al lui Poincaré

5. Geometria eliptică sau riemanniană

În geometria euclidiană, ca și în geometria hiperbolică sau a lui Bolyai-Lobacevski, se face ipoteza tacită că dreapta este infinită (întinderea infinită a dreptei este esențial legată de conceptul și de axiomele relației „între”). Dar după ce geometria hiperbolică a deschis drumul către construirea liberă a geometriilor, a fost natural ca matematicienii să se întrebe dacă nu se pot construi alte geometrii neeuclidiene, în care o dreaptă să nu fie infinită, ci finită și închisă. Desigur, în astfel de geometrii nu numai postulatul paralelelor, dar și axiomele legate de relația „între” trebuie să fie părăsite. Cercetări moderne au scos în evidență importanța fizică a acestor geometrii. Ele au fost considerate pentru prima dată în disertația inaugurală susținută în 1851 de Riemann, cu prilejul admiterii sale ca privat-docent al Universității din Göttingen. Geometriile cu drepte finite închise pot fi construite fără nici o greutate. Să ne închipuim o lume bidimensională, formată din suprafața S a unei sfere, în care definim „dreapta” ca fiind un cerc mare al sferei. Acesta ar fi modul natural de descriere a lumii unui navigator, deoarece arcele de cercuri mari sînt curbele de lungime minimă între două puncte de pe o sferă, iar aceasta este proprietatea caracteristică a dreptelor din plan. Într-o astfel de lume, *oricare* două „drepte” se intersectează, astfel încît dintr-un punct exterior *nu* se poate

duce nici o paralelă (dreaptă care nu intersectează) la o dreaptă dată. Geometria „dreptelor" din această lume se numește *geometrie eliptică*. În această geometrie, distanța dintre două puncte se măsoară prin lungimea celui mai scurt arc al cercului mare care unește punctele. Unghiurile se măsoară ca în geometria euclidiană. În general, considerăm ca fiind caracteristic pentru geometria eliptică faptul că nu există paralele la o dreaptă.

Urmînd calea propusă de Riemann, putem generaliza această geometrie în felul următor. Să considerăm o lume care constă dintr-o suprafață curbată din spațiu, nu neapărat sferică, și să definim „dreapta" care unește două puncte, ca fiind *curba de lungime minimă* sau „geodezica" care unește aceste puncte. Punctele de pe suprafață pot fi împărțite în două clase: 1. Puncte în vecinătatea cărora suprafața este ca o sferă, în sensul că ea se află în întregime de o parte a planului tangent în acel punct. 2. Puncte în vecinătatea cărora suprafața are forma unei șei și se află de ambele părți ale planului tangent în punctul considerat. Punctele de primul fel se numesc puncte eliptice ale suprafeței, deoarece dacă planul tangent este deplasat puțin, paralel cu el însuși, el intersectează suprafața după o curbă eliptică; punctele de al doilea fel se numesc hiperbolice, deoarece dacă planul tangent este deplasat puțin, paralel cu el în-

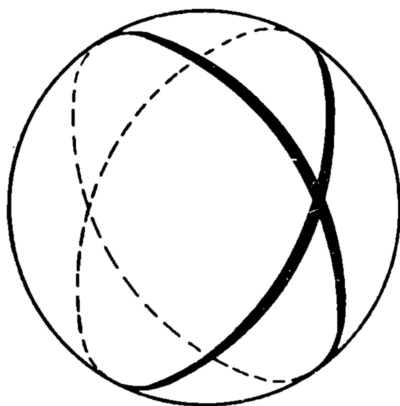


Fig. 113. „Drepte" pentru o geometrie riemanniană

suși, el intersectează suprafața după o curbă asemănătoare cu hiperbola. Geometria „dreptelor" geodezice din vecinătatea unui punct al suprafeței este eliptică sau hiperbolică, după cum punctul este eliptic sau hiperbolic. Într-un astfel de model de geometrie neeuclidiană, unghiurile se măsoară ca în geometria euclidiană.

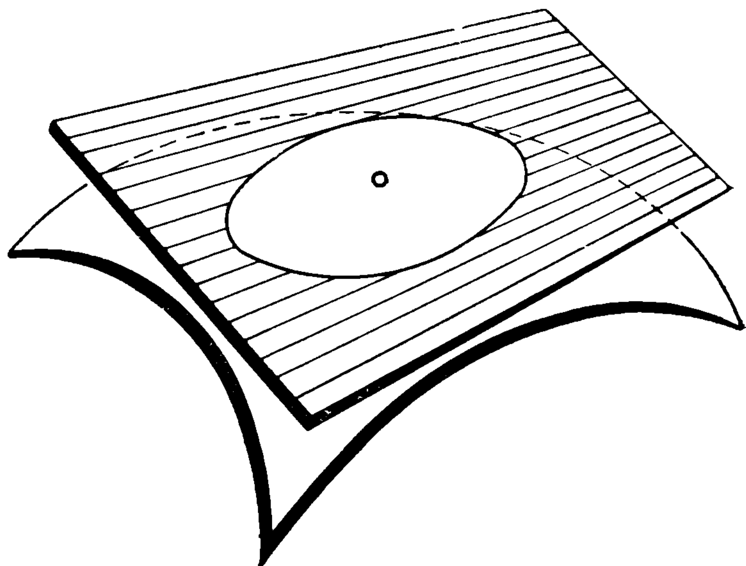


Fig. 114. Puncte eliptice

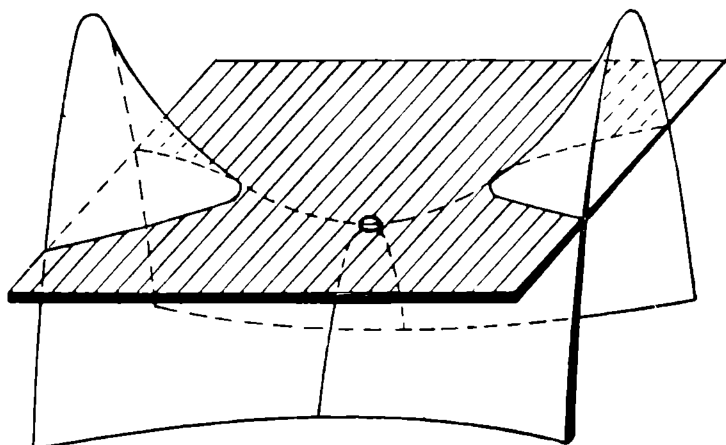


Fig. 115. Puncte hiperbolice

Această idee a fost dezvoltată de Riemann, care a considerat o geometrie a spațiului, analogă cu această geometrie a suprafeței, în care „curbura"spațiului poate schimba natura geometriei de la punct la punct. „Dreptele" unei geometrii riemanniene sînt geodezicele. În teoria generală a relativității a lui Einstein, geometria spațiului este o geometrie riemanniană, lumina se propagă în lungul geodezicelor, iar curbura spațiului este determinată de natura materiei pe care o conține.

De la originea sa, aflată în studiul axiomaticii, geometria neeuclidiană a devenit un instrument extrem de util pentru aplicații la lumea fizică. În teoria relativității, în optică și în teoria generală a propagării undelor, o descriere neeuclidiană a fenomenelor este uneori mult mai adecvată decît una euclidiană.

APENDICE

GEOMETRIE ÎN SPAȚII CU MAI MULT DE TREI DIMENSIUNI

1. Introducere

„Spațiul real", adică mediul experienței noastre fizice, are trei dimensiuni, planul are două dimensiuni, iar dreapta numai una. Intuiția noastră spațială, în sensul obișnuit al cuvîntului, este strict limitată la trei dimensiuni. Totuși, în multe împrejurări, este comod să vorbim despre „spații" care au patru sau mai multe dimensiuni. Ce se înțelege prin spațiu n-dimensional, dacă n este mai mare decît 3, și la ce scopuri poate să servească? Se poate da un răspuns atît din punct de vedere analitic, cît și din punct de vedere pur geometric. Terminologia spațiului n-dimensional poate fi privită doar ca limbaj geometric sugestiv pentru idei matematice, care nu mai sînt accesibile intuiției geometrice obișnuite. Vom indica pe scurt considerațiile simple, care motivează și justifică acest limbaj.

2. Abordarea analitică

Am remarcat deja schimbarea rolului geometriei analitice, petrecută în cursul dezvoltării ei. Punctele, dreptele, curbele etc. erau considerate la început ca entități pur „geometrice", iar sarcina geometriei analitice era doar să le asocieze sisteme de numere sau ecuații și să interpreteze sau să dezvolte teoria geometrică prin metode algebrice sau analitice. Cu trecerea timpului, a început să se impună din ce în ce mai mult punctul de vedere opus. Un număr x, sau o pereche de numere x, y, sau un triplet de numere x, y, z erau considerate ca obiecte fundamentale, iar apoi aceste entități analitice erau „interpretate" ca

puncte pe o dreaptă, într-un plan sau în spațiu. Din acest punct de vedere, limbajul geometric servește doar la enunțarea relațiilor existente între numere. Putem renunța la caracterul primar sau chiar independent al obiectelor geometrice, spunînd că o pereche de numere x, y *este* un punct din plan, mulțimea tuturor perechilor de numere x, y care satisfac ecuația liniară $L(x, y) = ax + by + c = 0$ (unde a, b, c sînt constante date), *este* o dreaptă etc. Definiții asemănătoare pot fi date în spațiul cu trei dimensiuni.

Chiar dacă ne interesează mai ales o problemă algebrică, se poate întîmpla ca limbajul geometriei să permită o descriere scurtă și adecvată a ei, și ca intuiția geometrică să sugereze procedeul algebric potrivit de rezolvare a ei. De exemplu, dacă dorim să rezolvăm un sistem de trei ecuații liniare, cu trei necunoscute x, y, z:

$$L(x, y, z) = ax + by + cz + d = 0$$

$$L'(x, y, z) = a'x + b'y + c'z + d' = 0$$

$$L''(x, y, z) = a''x + b''y + c''z + d'' = 0,$$

putem interpreta geometrie problema sub forma găsirii punctului de intersecție, în spațiul tridimensional R_3, a trei plane, definite de ecuațiile $L = 0$, $L' = 0$, $L'' = 0$. Apoi, dacă considerăm doar perechile de numere x, y pentru care $x > 0$, le putem interpreta geometric sub forma semiplanului aflat la dreapta axei Ox. Mai general, mulțimea tuturor perechilor de numere x, y pentru care

$$L(x, y) = ax + by + c > 0$$

poate fi interpretată geometric sub forma unui semiplan, aflat de o parte a dreptei $L = 0$, iar mulțimea tripletelor de numere x, y, z pentru care

$$L(x, y, z) = ax + by + cz + d > 0$$

poate fi interpretată geometric sub forma unui semispațiu, aflat de o parte a planului $L(x, y, z) = 0$.

Introducerea unui „spațiu cvadridimensional" sau chiar a unui „spațiu n-dimensional" este acum foarte naturală. Să considerăm un cvadruplet de numere x, y, z, t. Un astfel de cvadruplet este reprezentat de, sau, mai simplu, *este* un punct din spațiul cu patru dimensiuni R_4. Mai general, un punct din spațiul n-dimensional R_n este, prin definiție, un sistem ordonat de n-numere reale $x_1, x_2, \ldots, x_n$. Nu are importanță faptul că nu ne putem imagina un astfel de punct. Limbajul geometric rămîne tot atît de sugestiv pentru proprietățile algebrice, în care intervin patru sau n variabile. Cauza acestui fapt constă în faptul că multe din proprietățile algebrice ale ecuațiilor liniare etc. sînt esențial independente de numărul de variabile care intervin sau, după cum se spune, de dimensiunea spațiului variabilelor. De exemplu, numim

„hiperplan" mulțimea tuturor punctelor $x_1, x_2, \ldots, x_n$ din spațiul n-dimensional R_n, care satisfac o ecuație liniară:

$$L(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n + b = 0.$$

Atunci problema algebrică fundamentală a rezolvării unui sistem de n-ecuații liniare, cu n necunoscute,

$$\begin{aligned} L_1(x_1,x_2, \ldots,x_n) &= 0 \\ L_2(x_1,x_2, \ldots,x_n) &= 0 \\ \ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots \\ L_n(x_1,x_2, \ldots,x_n) &= 0, \end{aligned}$$

se enunță în limbaj geometric sub forma găsirii punctului de intersecție a n hiperplane: $L_1 = 0$, $L_2 = 0$, $\ldots$, $L_n = 0$.

Avantajul acestui mod geometric de exprimare constă doar în faptul că el scoate în evidență anumite aspecte algebrice care sînt independente de n *și care pot fi reprezentate geometric dacă* $n \leqslant 3$. În numeroase aplicații, folosirea unei astfel de terminologii prezintă avantajul prescurtării, facilitării și dirijării considerațiilor analitice intrinseci. Teoria relativității poate fi menționată ca un exemplu în care s-a obținut un progres important, reunind coordonatele spațiale x, y, z ale unui „eveniment", cu coordonata temporală t, într-o varietate „spațiu-timp" cu patru dimensiuni a cvadrupletelor x, y, z, t. Prin introducerea unei geometrii hiperbolice neeuclidiene în acest cadru analitic a devenit posibilă descrierea deosebit de simplă a multor situații complexe. Avantaje asemănătoare au fost obținute în mecanică și în fizica statistică, ca și în domenii pur matematice.

Iată cîteva exemple din matematică. Mulțimea tuturor cercurilor din plan formează o varietate tridimensională, deoarece un cerc cu centrul x, y și raza t poate fi reprezentat printr-un punct cu coordonatele x,y,t. Deoarece raza unui cerc este un număr pozitiv, mulțimea punctelor care reprezintă cercuri umple un semispațiu. În același mod, mulțimea tuturor sferelor din spațiul tridimensional obișnuit formează o varietate cu patru dimensiuni, deoarece orice sferă de centru x, y, z și de rază t poate fi reprezentată printr-un punct cu coordonatele x, y, z, t. Un cub din spațiul tridimensional, cu muchia de lungime 2, cu fețele paralele cu planele de coordonate și cu centrul în origine constă din mulțimea tuturor punctelor x_1, x_2, x_3 pentru care $|x_1| \leqslant 1$, $|x_2| \leqslant 1$, $|x_3| \leqslant 1$. În același mod, un „cub" din spațiul n-dimensional R_n, cu muchia egală cu 2, cu fețele paralele cu planele de coordonate și cu centrul în origine, este definit ca fiind totalitatea punctelor $x_1, x_2, \ldots, x_n$ pentru care sînt simultan îndeplinite inegalitățile

$$|x_1| \leqslant 1, \quad |x_2| \leqslant 1, \ \ldots, \ |x_n| \leqslant 1.$$

„Suprafața" acestui cub constă din toate punctele cubului pentru care are loc cel puțin o egalitate. Elementele de suprafață de dimensiune $n - 2$ constau din acele puncte în care au loc cel puțin *două* egalități etc.

Exercițiu : Descrieți suprafața unui astfel de cub, în cazurile tri-, cvadri- și n-dimensionale.

*3. Abordarea geometrică sau combinatorie

În timp ce abordarea analitică a geometriei n-dimensionale este simplă și bine adaptată celor mai multe aplicații, mai există o metodă de tratare, de natură pur geometrică. Ea se bazează pe reducerea de la datele n-dimensionale la cele $(n - 1)$-dimensionale, care ne permite să definim geometria în spații cu mai multe dimensiuni, prin inducție matematică.

Să începem cu frontiera unui triunghi ABC în spațiul cu două dimensiuni. Secționînd poligonul închis în punctul C și rotind apoi pe AC și pe BC, așternîndu-le în lungul dreptei AB, obținem o figură dreaptă simplă, ca în fig. 116, în care punctul C apare de două ori. Această figură unidimensională dă o reprezentare completă a frontierei triunghiului bidimensional. Îndoind segmentele AC și BC în plan, putem face ca cele două puncte C să coincidă din nou. Dar, și aici este momentul important, nu este necesar să facem această îndoire. Trebuie numai să convenim să „identificăm", adică să nu facem distincție, între cele două puncte C din fig. 116, chiar dacă ele nu coincid efectiv, ca entități geometrice în sensul naiv al cuvîntului. Apoi putem face încă un pas ; separînd cele trei segmente în punctele A și B, obținem un sistem

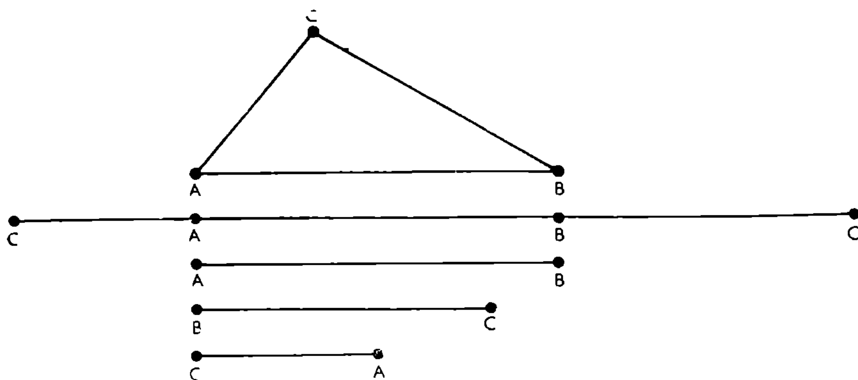


Fig. 116. Triunghiul definit prin segmente cu capetele identificate

de trei segmente CA, AB, BC, care pot fi reunite din nou, astfel încît să formeze un triunghi „real", iar perechile identificate de puncte să coincidă între ple. Această idee a identificării diferitelor puncte dintr-un sistem de segmente, eentru a forma un poligon (în cazul nostru un triunghi), este adesea foarte

utilă. Dacă dorim să confecționăm un cadru complicat din bare de oțel, ca de pildă cadrul unui pod, îl confecționăm din bare simple și notăm cu același simbol capetele care trebuie îmbinate atunci cînd cadrul trebuie montat în spațiu. Sistemul de bare cu capete marcate este pe deplin echivalent cu structura spațială. Această remarcă sugerează calea pe care putem reduce un poliedru

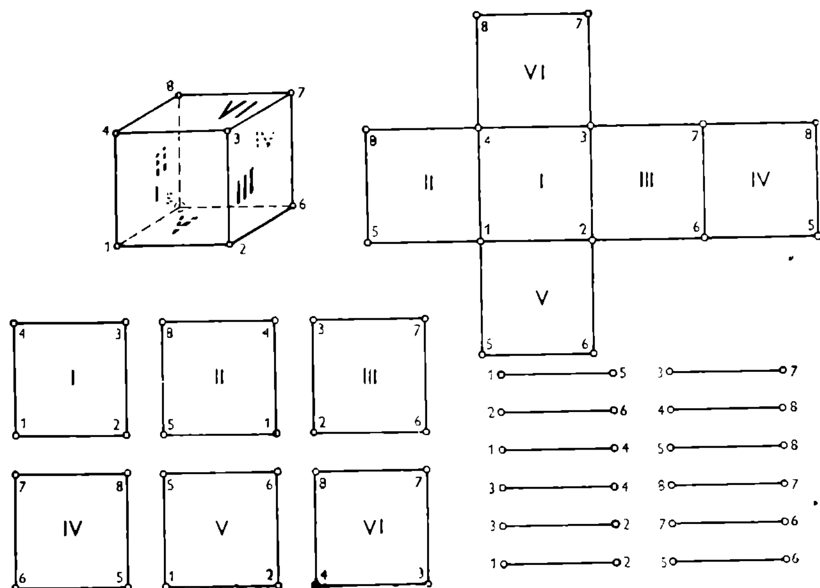


Fig. 117. Cubul definit prin marcarea vîrfurilor și a muchiilor

bidimensional din spațiul tridimensional la figuri de dimensiuni mai mici. Să luăm, de pildă, suprafața unui cub (fig. 117). Ea poate fi redusă imediat la un sistem de șase pătrate plane, ale căror segmente frontieră trebuie identificate în mod potrivit, iar al doilea pas constă în considerarea unui sistem de 12 segmente, cu extremități identificate în mod corespunzător.

În general, orice poliedru din spațiul tridimensional R_3 poate fi redus în acest mod fie la un sistem de poligoane plane, fie la un sistem de segmente rectilinii.

[*Exercițiu*: Efectuați această reducere pentru toate poliedrele regulate (cf. p. 254).

Acum este foarte clar că putem inversa raționamentul nostru, *definind* un poligon din plan printr-un sistem de segmente rectilinii, iar un poliedru din R_3 printr-un sistem de poligoane din R_2 sau, printr-o nouă reducere, cu ajutorul unui sistem de segmente rectilinii. Deci este natural să definim un „poliedru"

din spațiul cu patru dimensiuni R_4, printr-un sistem de poliedre din R_3, cu o identificare potrivită a fețelor sale bidimensionale; poliedrele din R_5, prin sisteme de poliedre din R_4, și așa mai departe. În cele din urmă, putem reduce orice poliedru din R_n la un sistem de segmente rectilinii.

Nu este posibil să dezvoltăm în continuare acest subiect. Să mai adăugăm doar cîteva observații, fără a face demonstrație. Un cub din R_4 este mărginit de opt cuburi tridimensionale, fiecare fiind identificat cu un ,,vecin'' de-a lungul unei fețe bidimensionale. Cubul din R_4 are 16 vîrfuri, iar în fiecare din acestea se întîlnesc patru din cele 32 de muchii. În R_4 există șase poliedre regulate. Pe lîngă ,,cub'', mai există un poliedru, mărginit de cinci tetraedre regulate, un poliedru mărginit de 16 tetraedre, un poliedru mărginit de 24 de octaedre, un poliedru mărginit de 120 dodecaedre și un poliedru mărginit de 600 de tetraedre. Pentru $n > 4$ dimensiuni, s-a demonstrat că sînt posibile doar trei poliedre regulate: un poliedru cu $n + 1$ vîrfuri, mărginit de $n + 1$ poliedre din R_{n-1} cu n fețe, de cîte $(n-2)$ dimensiuni; un poliedru cu 2^n vîrfuri, mărginit de $2n$ poliedre din R_{n-1}, cu $2n - 2$ fețe; și un poliedru cu $2n$ vîrfuri și 2^n poliedre cu n fețe din R_{n-1}, ca frontiere.

Exercițiu: Comparați definiția cubului din R_4, dată în secțiunea 2, cu definiția dată în această secțiune și arătați că definiția ,,analitică'' a suprafeței cubului din secțiunea 2 este echivalentă cu definiția ,,combinatorie'' din această secțiune.

Din punct de vedere structural sau ,,combinatoriu'', cele mai simple figuri geometrice de dimensiuni 0, 1, 2, 3 sînt respectiv punctul, segmentul, triunghiul și tetraedrul. Pentru a avea o notație uniformă, să notăm aceste figuri respectiv

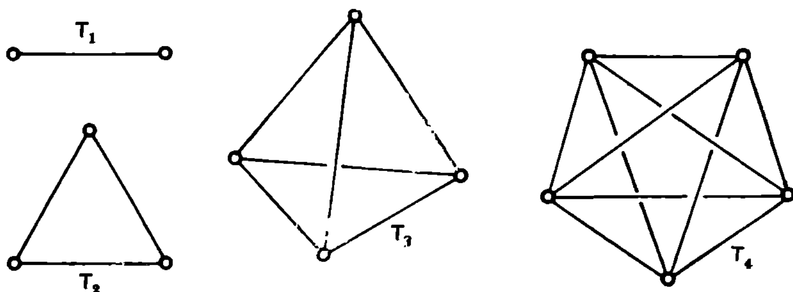


Fig. 118. Cele mai simple elemente cu 1, 2, 3 și 4 dimensiuni

prin simbolurile T_0, T_1, T_2, T_3 (indicele arată dimensiunea). Structura fiecăreia din aceste figuri este descrisă prin afirmația că fiecare T_n conține $n + 1$ vîrfuri și că fiecare submulțime formată din $i + 1$ vîrfuri ale lui T_n $(i = 0, 1, \ldots, n)$ determină un T_i. De exemplu, tetraedrul tridimensional T_3 conține 4 vîrfuri, 6 muchii și 4 triunghiuri.

Acum este limpede cum trebuie să continuăm. Definim un „tetraedru" patru-dimensional T_4 ca fiind o mulțime de cinci vîrfuri, astfel încît orice submulțime formată din patru vîrfuri determină un T_3, fiecare submulțime de trei vîrfuri determină un T_2 etc. Diagrama schematică a lui T_4 este arătată în fig. 118; vedem că T_4 conține 5 vîrfuri, 10 muchii, 10 triunghiuri și 5 tetraedre.

Generalizarea la n dimensiuni este imediată. Din teoria combinărilor se știe că există $C_r^i = \dfrac{r!}{i!(r-i)!}$ submulțimi diferite formate din i obiecte, luate dintr-o mulțime dată, formată din r obiecte. Deci un „tetraedru" n-dimensional conține

$$C_{n+1}^1 = n + 1 \qquad \text{vîrfuri } (T_0),$$

$$C_{n+1}^2 = \frac{(n+1)!}{2!(n-1)!} \qquad \text{segmente } (T_1),$$

$$C_{n+1}^3 = \frac{(n+1)!}{3!(n-2)!} \qquad \text{triunghiuri } (T_2),$$

$$C_{n+1}^4 = \frac{(n+1)!}{4!(n-3)!} \qquad T_3\text{-uri}.$$

. . . .

$$C_{n+1}^{n+1} = 1 \qquad T_n\text{-uri}.$$

Exercițiu: Trasați diagrama lui T_5 și determinați numărul diferiților T_i-uri pe care îi conține, pentru $i = 0, 1, \ldots, 5$.

Capitolul V

TOPOLOGIA

INTRODUCERE

La mijlocul secolului al XIX-lea a început o dezvoltare cu totul nouă a geometriei, care curînd a devenit una dintre marile forțe ale matematicii moderne. Noul subiect, denumit topologie sau *analysis situs*, are ca obiect studiul proprietăților figurilor geometrice care persistă, chiar dacă figurile sînt supuse unor transformări care distrug toate proprietățile lor metrice și proiective.

Unul dintre marii geometri ai timpului a fost A. F. Moebius (1790—1868), un om pe care modestia l-a destinat unei cariere de astronom neînsemnat într-un observator german de mîna a doua. La vîrsta de 68 de ani el a supus Academiei din Paris un memoriu, referitor la suprafețele „unilatere", care conținea cîteva din cele mai surprinzătoare fapte ale noii ramuri a geometriei. Ca și alte contribuții importante de mai înainte, lucrarea lui a rămas înmormîntată ani de zile în dosarele Academiei, pînă ce a fost publicată de autor. Independent de Moebius, astronomul J. B. Listing (1808—1882) din Göttingen, a făcut descoperiri asemănătoare și, la sugestia lui Gauss, a publicat în 1847 o cărticică cu titlul *Vorstudien zur Topologie*. Cînd Bernhard Riemann (1826—1866) a venit la Göttingen ca student, atmosfera matematică a acestui orășel universitar era plină de interes pentru aceste noi și ciudate idei geometrice. Curînd el și-a dat seama că în aceste idei se află cheia înțelegerii celor mai profunde proprietăți ale funcțiilor analitice de o variabilă complexă. Poate că nimic nu a dat un impuls mai puternic dezvoltării ulterioare a topologiei decît marea construcție a teoriei lui Riemann a funcțiilor, în care noțiunile topologice sînt absolut fundamentale.

La început, noutatea metodelor din noul domeniu de studiu nu a lăsat matematicienilor timpul necesar pentru a prezenta rezultatele lor sub forma axiomatică tradițională, a geometriei elementare. În schimb, pionieri, ca de pildă Poincaré, au fost siliți să se bazeze în mare măsură pe intuiția geometrică. Chiar și astăzi, cel care studiază topologia găsește că, insistînd prea mult asupra unei forme riguroase de prezentare, ar putea pierde din aspectul conți-

nutului geometric esențial, datorită numeroaselor amănunte formale. Totuşi un mare merit al lucrărilor recente constă în încadrarea topologiei într-un sistem matematic riguros, în care intuiția rămîne izvorul, dar nu şi confirmarea finală a adevărului. În timpul acestui proces de dezvoltare, început de L. E. J. Brouwer, importanța topologiei pentru întreaga matematică a crescut mereu. Matematicieni americani, ca de pildă O. Veblen, J. W. Alexander şi S. Lefschetz, au adus contribuții importante acestei teorii.

În timp ce topologia este în mod cert o creație a ultimei sute de ani, au existat cîteva descoperiri premergătoare izolate, care mai tîrziu şi-au găsit locul în dezvoltarea sistematică modernă. Cea mai importantă dintre ele este o formulă, care leagă numărul de vîrfuri, de muchii şi de fețe ale unui poliedru simplu, observată încă în 1640 de Descartes şi redescoperită şi folosită de Euler în 1752. Natura esențială a acestei relații, ca teoremă topologică, a devenit evidentă mult mai tîrziu, după ce Poincaré a recunoscut „formula lui Euler" şi generalizările ei, ca fiind una dintre teoremele centrale ale topologiei. Astfel, din motive atît istorice, cît şi intrinseci, vom începe discuția noastră asupra topologiei cu formula lui Euler. Deoarece idealul unei rigori perfecte nu este nici necesar şi nici de dorit la primii paşi făcuți într-un domeniu necunoscut, nu vom ezita să apelăm uneori la intuiția geometrică a cititorului.

§ 1. FORMULA LUI EULER PENTRU POLIEDRE

Cu toate că studiul poliedrelor a ocupat o poziție centrală în geometria greacă, a fost meritul lui Descartes şi al lui Euler de a descoperi următorul fapt: Într-un poliedru simplu, fie V—numărul vîrfurilor, M—numărul muchiilor şi F—numărul fețelor; atunci avem întotdeauna

$$V - M + F = 2. \qquad (1)$$

Prin *poliedru* se înțelege un solid, a cărui suprafață constă dintr-un număr de fețe poligonale. În cazul solidelor regulate, toate poligoanele sînt congruente şi toate unghiurile din vîrfuri sînt egale. Un poliedru este *simplu*, dacă nu are „găuri", astfel încît suprafața sa poate fi deformată în mod continuu, în suprafața unei sfere. Fig. 120 arată un poliedru simplu, care nu este regulat, în timp ce fig. 121 arată un poliedru care nu este simplu.

Cititorul ar trebui să verifice faptul că formula lui Euler este valabilă pentru poliedrele simple din fig. 119 şi 120, dar nu rămîne în vigoare pentru poliedrul din fig. 121.

Pentru a demonstra formula lui Euler, să ne închipuim că poliedrul simplu dat este gol pe dinăuntru, iar suprafața lui e confecționată din foiță de cauciuc subțire. Atunci, dacă îndepărtăm una dintre fețele poliedrului, putem deforma suprafața rămasă, întinzînd-o pe un plan. Desigur, ariile fețelor şi unghiurile dintre muchiile poliedrului se vor schimba prin această operație, însă rețeaua

de vîrfuri şi muchii din plan va conţine acelaşi număr de vîrfuri şi de muchii ca şi reţeaua poliedrului iniţial, în timp ce numărul poligoanelor va fi micşorat cu o unitate faţă de cel al poliedrului iniţial, deoarece am îndepărtat o faţă. Vom arăta acum că pentru reţeaua plană avem $V - M + F = 1$, astfel încît

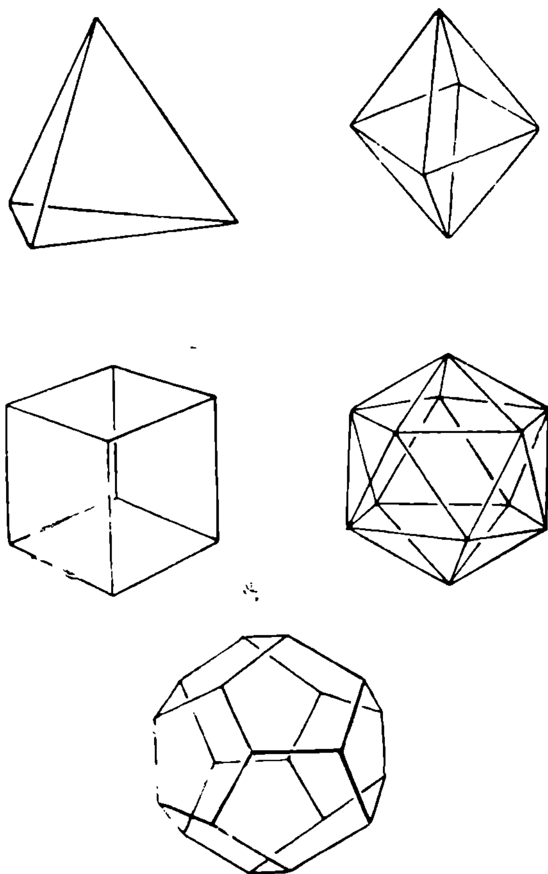


Fig. 119. Poliedrele regulate

dacă ţinem seama de faţa îndepărtată, rezultatul pentru poliedrul iniţial va fi $V - M + F = 2$.

Mai întîi „triangulăm" reţeaua plană în modul următor: Într-un poligon al reţelei care nu este triunghi, să ducem o diagonală. În acest mod, facem pe M şi pe F să crească cu 1, ceea ce nu schimbă valoarea lui $V - M + F$. Ducem,

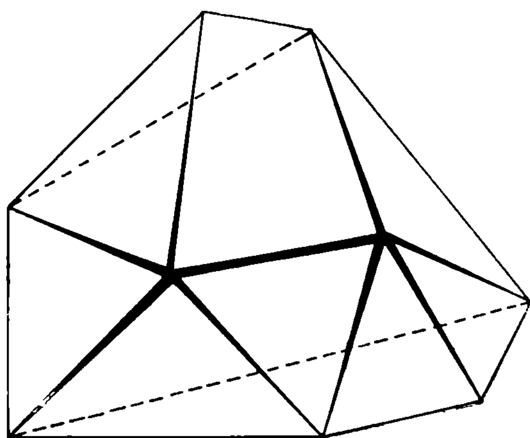


Fig. 120. Poliedru simplu; **V — M + + F = 9 — 18 + 11 = 2**

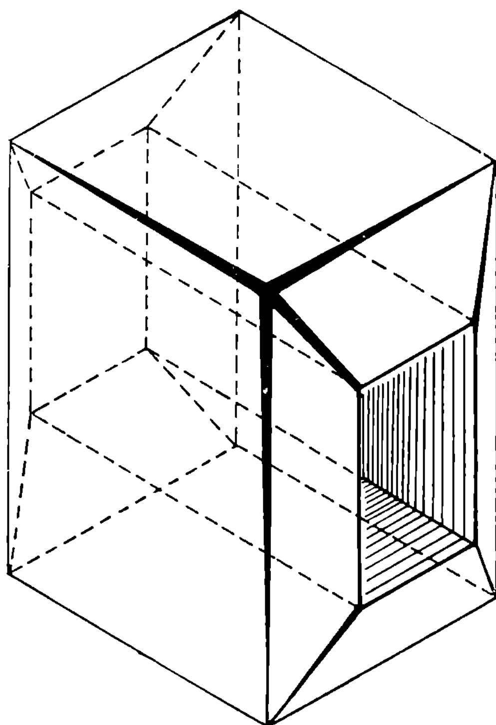


Fig. 121. Poliedru ne-simplu; **V — M + F = 16 — 32 + 16 = 0.**

în continuare, diagonale, care unesc perechi de puncte (fig. 122) pînă ce figura este formată numai din triunghiuri. În rețeaua triangulată, $V - M + F$ are aceeași valoare ca și înaintea triangulării, deoarece ducînd diagonalele, nu am modificat-o. Unele triunghiuri vor avea laturile pe frontiera rețelei plane. Dintre acestea, unele, ca de pildă ABC, vor avea o singură latură pe frontieră,

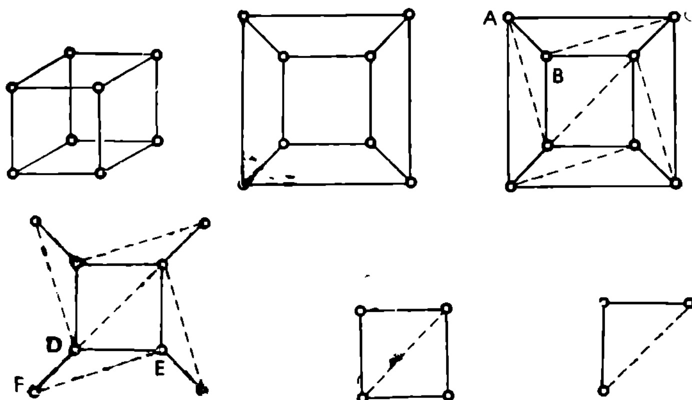


Fig. 122. Demonstrarea teoremei lui Euler

în timp ce alte triunghiuri pot avea două laturi pe frontieră. Luăm un triunghi frontieră arbitrar și îndepărtăm acea parte care nu mai aparține unui alt triunghi. Astfel, din ABC îndepărtăm latura AC și fața, lăsînd vîrfurile A, B, C și muchiile AB, BC; din DEF îndepărtăm fața, muchiile DF, FE și vîrful F. Îndepărtarea unui triunghi de tipul ABC micșorează pe M și pe F cu o unitate, în timp ce V rămîne neschimbat, astfel încît $V - M + F$ rămîne neschimbat. Îndepărtarea unui triunghi de tipul DEF micșorează pe V cu o unitate, pe M cu două unități, iar pe F cu o unitate, astfel încît $V - M + F$ rămîne iar neschimbat. Alegînd în mod potrivit un șir de operații de acest fel, putem îndepărta triunghiurile care au laturile pe frontieră (care se modifică la fiecare operație) pînă ce, în cele din urmă, rămîne un singur triunghi, cu cele trei laturi, trei vîrfuri și o față. Pentru această rețea simplă, avem $V - M + F = 3 - 3 + 1 = 1$. Însă am văzut că, ștergînd mereu cîte un triunghi, $V - M + F$ nu se modifică. De aceea, $V - M + F$ trebuie să fie egal cu 1 atît pentru rețeaua plană inițială, cît și pentru poliedrul din care lipsește o față. Deducem că $V - M + F = 2$ pentru poliedrul complet. Aceasta completează demonstrația formulei lui Euler (cf. (56), (57), p. 516).

Pe baza formulei lui Euler, este ușor de arătat că există cel mult cinci poliedre regulate. Într-adevăr, să presupunem că un poliedru regulat are F fețe, fiecare față fiind un poligon regulat

cu n laturi şi că r muchii se întîlnesc în fiecare vîrf. Numărînd muchiile după numărul fețelor și al vîrfurilor, vedem că

$$(2) \qquad nF = 2M;$$

deoarece fiecare muchie aparţine la două feţe şi deci este numărată de două ori în produsul nF; mai mult

$$(3) \qquad rV = 2M,$$

deoarece fiecare muchie are două vîrfuri. Deci din (1) obţinem egalitatea

$$\frac{2M}{n} + \frac{2M}{r} - M = 2$$

sau

$$(4) \qquad \frac{1}{n} + \frac{1}{r} = \frac{1}{2} + \frac{1}{M}.$$

Observăm mai întîi că $n \geqslant 3$ şi $r \geqslant 3$, deoarece un poligon trebuie să aibă cel puţin trei laturi şi cel puţin trei muchii trebuie să pornească din fiecare vîrf al poliedrului. Însă n şi r nu pot fi amîndouă *mai mari* decît 3, deoarece atunci membrul stîng al egalităţii (4) nu ar putea depăşi pe 1/2, ceea ce este imposibil, oricare ar fi valoarea pozitivă a lui M. De aceea, să vedem ce valori poate avea r dacă $n = 3$, şi ce valori poate avea n dacă $r = 3$. Mulţimea poliedrelor date de aceste două cazuri reprezintă numărul de poliedre regulate posibile.

Pentru $n = 3$, ecuaţia (4) devine

$$\frac{1}{r} - \frac{1}{6} = \frac{1}{M};$$

astfel încît r poate fi egal cu 3, 4 sau 5 (6, sau orice număr mai mare trebuie exclus, deoarece $1/M$ este întotdeauna pozitiv). Pentru aceste valori ale lui n şi r găsim că $M = 6$, 12 sau 30, corespunzînd respectiv tetraedrului, octaedrului şi icosaedrului. În acelaşi mod, pentru $r = 3$, obţinem egalitatea

$$\frac{1}{n} - \frac{1}{6} = \frac{1}{M},$$

din care rezultă că $n = 3$, 4 sau 5 şi $M = 6$, 12 sau 30. Aceste valori corespund respectiv tetraedrului, cubului şi dodecaedrului. Substituind aceste valori ale lui n, r şi M în egalităţile (2) şi (3) obţinem numărul de vîrfuri şi feţe ale poliedrelor corespunzătoare.

§ 2. PROPRIETĂŢILE TOPOLOGICE ALE FIGURILOR

1. Proprietăţi topologice

Am arătat că formula lui Euler este valabilă pentru orice poliedru simplu, însă domeniul de valabilitate al acestei formule depăşeşte cu mult poliedrele geometriei elementare, cu feţele lor plane şi muchiile lor rectilinii. Demon-

strația pe care am dat-o s-ar aplica tot atît de bine unui poliedru simplu, cu fețe și muchii curbate, sau oricărei subdiviziuni a suprafeței unei sfere în regiuni mărginite de arce curbe. Mai mult, dacă ne imaginăm că suprafața poliedrului sau a sferei este confecționată dintr-o foaie subțire de cauciuc, formula lui Euler și-ar păstra valabilitatea dacă suprafața este deformată prin îndoirea și întinderea cauciucului în orice mod, atîta timp cît nu au loc rupturi. Într-adevăr, formula se referă doar la *numerele* vîrfurilor, muchiilor și fețelor și nu la lungimi, arii, coliniaritate, birapoarte sau oricare din noțiunile obișnuite ale geometriei elementare sau proiective.

Reamintim că geometria elementară se ocupă cu mărimile (lungime, unghi și arie) care rămîn neschimbate prin deplasări rigide, în timp ce geometria proiectivă se ocupă de noțiuni (punct, dreaptă, incidență și biraport), care rămîn neschimbate față de grupul mai cuprinzător al transformărilor proiective. Însă mișcările rigide și proiecțiile sînt cazuri particulare de *transformări topologice*: o transformare topologică a unei figuri geometrice A într-o altă figură A', este dată de orice corespondență

$$p \leftrightarrow p'$$

între punctele p ale lui A și punctele p' ale lui A', care are următoarele două proprietăți:

1. *Corespondența este biunivocă.* Aceasta înseamnă că fiecărui punct p al lui A îi corespunde un singur punct p' al lui A', și reciproc.

2. *Corespondența este continuă în ambele sensuri.* Aceasta înseamnă că dacă luăm două puncte oarecare p, q ale lui A și-l deplasăm pe p, astfel încît distanța dintre el și q să tindă spre zero, atunci distanța dintre punctele corespunzătoare p', q' ale lui A' tinde, de asemenea, spre zero, și reciproc.

Orice proprietate a unei figuri geometrice A, care se menține pentru orice figură în care A poate fi transformată printr-o transformare topologică, se numește *proprietate topologică* a lui A, iar *topologia* este ramura geometriei care se ocupă numai de proprietățile topologice ale figurilor. Să ne imaginăm o figură copiată „liber" de către un desenator conștiincios, dar nepriceput, care face dreptele curbe și modifică unghiurile, distanțele și ariile; atunci, cu toate că proprietățile metrice și proiective ale figurii inițiale se vor pierde, proprietățile topologice rămîn neschimbate.

Cele mai intuitive exemple de transformări topologice generale sînt *deformările*. Imaginați-vă o figură, ca de pildă o sferă sau un triunghi, confecționată sau desenată pe o foaie subțire de cauciuc, care este apoi întinsă și îndoită oricum, fără a o rupe și fără a face să coincidă puncte diferite (dacă am face să coincidă puncte diferite am încălca condiția 1, iar ruperea foii de cauciuc ar încălca condiția 2, deoarece două puncte ale figurii inițiale care tind să coincidă din părți opuse ale curbei în lungul căreia a fost ruptă foaia, nu ar tinde să coincidă în figura ruptă). Poziția finală a figurii va fi atunci o imagine topologică a celei inițiale. Un triunghi poate fi deformat în orice alt

triunghi sau într-un cerc, sau o elipsă, şi deci aceste figuri au aceleaşi proprietăţi topologice. Însă nu putem deforma un cerc într-un segment rectiliniu şi nici nu putem deforma suprafaţa unei sfere în suprafaţa laterală a unui cilindru.

Noţiunea generală de transformare topologică este mai largă decît noţiunea de deformare. De exemplu, dacă în cursul unei deformări tăiem o figură şi dacă coasem muchiile tăieturii după deformare în acelaşi mod ca şi mai înainte,



Fig. 123. Suprafeţe topologic echivalente



Fig. 124. Suprafeţe topologic neechivalente

procesul este tot o transformare topologică a figurii iniţiale, cu toate că el nu este o deformare. Astfel, cele două curbe din fig. 134 (p.274) sînt topologic echivalente între ele, iar ambele sînt echivalente cu un cerc, deoarece ele pot fi tăiate, deznodate şi recusute. Dar este imposibil să deformăm una dintre curbe în cealaltă sau într-un cerc, fără a tăia mai întîi curba.

Proprietăţile topologice ale figurilor (aşa cum sînt date de teorema lui Euler sau de alte teoreme care vor fi discutate în acest paragraf) sînt de cel mai mare interes şi de cea mai mare importanţă în multe cercetări matematice. Dintr-un anumit punct de vedere, ele sînt cele mai profunde şi fundamentale dintre toate proprietăţile geometrice, deoarece ele persistă şi după cele mai drastice transformări.

2. Conexiune

Ca un alt exemplu de figuri care nu sînt topologic echivalente, putem considera domeniile plane din fig. 125. Primul dintre ele constă din toate punctele interioare unui cerc, în timp ce al doilea constă din toate punctele aflate între două cercuri concentrice. Orice curbă închisă care se află în domeniul *a* poate fi deformată în mod continuu, sau „contractată” într-un singur punct, rămînînd

în interiorul domeniului. Un domeniu care are această proprietate se numeşte *simplu conex*. Domeniul *b* nu este simplu conex. De pildă, un cerc concentric cu cele două cercuri frontieră şi aflat între ele nu poate fi contractat într-un singur punct, rămînînd în interiorul domeniului, deoarece în cursul acestui proces curba ar trebui să treacă prin centrul cercurilor, care nu este un punct al domeniului. Un domeniu care nu este simplu conex se numeşte *multiplu conex*. Dacă domeniul multiplu conex *b* este tăiat în lungul unei raze, ca în fig. 126, domeniul care rezultă este simplu conex.

Mai general, putem construi domenii cu două, trei sau mai multe „găuri", ca de pildă, domeniul din fig. 127. Pentru a transforma acest domeniu într-un domeniu simplu conex, sînt necesare două tăieturi. Dacă sînt necesare n—1 tăieturi care nu se intersectează, de la frontieră la frontieră, pentru a transforma

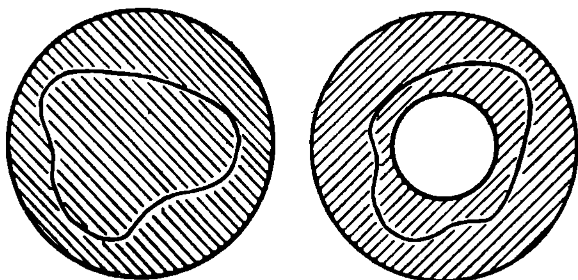


Fig. 125. Conexiunea simplă şi dublă

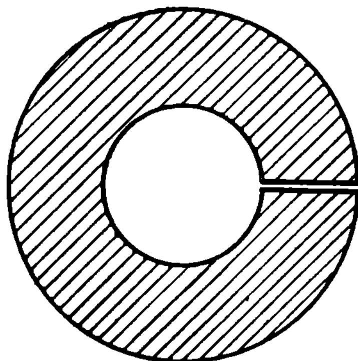


Fig. 126. Secţionarea unui domeniu dublu conex pentru a-l face simplu conex

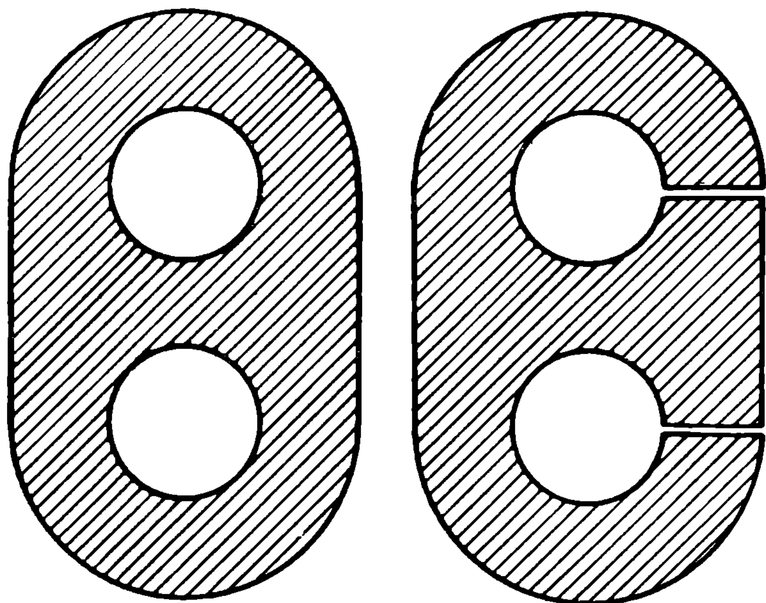


Fig. 127. Reducerea unui domeniu triplu conex

un domeniu multiplu conex dat D într-un domeniu simplu conex, domeniul D se numește n-conex. Ordinul de conexiune al unui domeniu plan este un ivariant topologic important al domeniului.

§ 3. ALTE EXEMPLE DE TEOREME TOPOLOGICE

1. Teorema lui Jordan

În plan este desenată o curbă simplă închisă (care nu se intersectează cu ea însăși). Care sînt proprietățile acestei figuri care se mențin, chiar dacă planul este privit ca foaie de cauciuc care poate fi deformată în orice mod? Lungimea curbei și aria pe care o cuprinde pot fi schimbate prin deformare. Există însă o proprietate topologică a configurației, care este atît de simplă, încît poate părea banală: *o curbă simplă închisă* C *din plan împarte planul în două domenii, unul interior și altul exterior.* Prin aceasta înțelegem că punctele planului pot fi grupate în două clase: A, exteriorul curbei și B interiorul ei, astfel încît orice pereche de puncte din aceeași clasă pot fi unite printr-o curbă care nu intersectează pe C, în timp ce orice curbă care unește o pereche de puncte care apar-

ţin la clase diferite trebuie să intersecteze pe *C*. Această propoziţie este evident adevărată pentru un cerc sau o elipsă, însă evidenţa dispare dacă privim o curbă mai complicată, ca de pildă poligonul din fig. 128.

Această teoremă a fost enunţată pentru prima dată de Camille Jordan (1838—1922) în celebrul său *Cours d'analyse* din care o întreagă generaţie de mate-

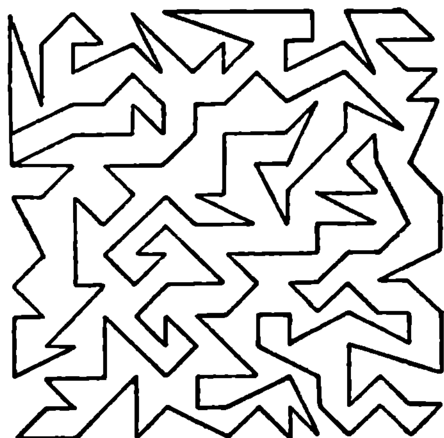


Fig. 128. Care puncte ale planului sînt interioare acestui poligon?

maticieni au învăţat noţiunea modernă de rigoare în analiza matematică. Destul de ciudat a fost faptul că demonstraţia dată de Jordan nu era nici scurtă, nici simplă, şi surpriza a fost cu atît mai mare, atunci cînd s-a arătat că demonstraţia lui Jordan nu era bună şi că este necesar un efort considerabil pentru a umple golurile raţionamentului. Primele demonstraţii riguroase ale teoremei erau foarte complicate şi greu de înţeles, chiar pentru mulţi matematicieni exersaţi. Doar de curînd au fost găsite demonstraţii comparativ simple. Unul dintre motivele dificultăţii se află în generalitatea noţiunii de ,,curbă simplă închisă", care nu se restrînge la clasa poligoanelor sau a curbelor ,,netede", ci include toate curbele care sînt imagini topologice ale cercului. Pe de altă parte, multe noţiuni ca de pildă ,,interior", ,,exterior" etc., care sînt atît de clare pentru intuiţie, trebuie să fie precizate în prealabil, înainte de a se putea ajunge la o demonstraţie riguroasă. Este de cea mai mare importanţă teoretică să se analizeze aceste noţiuni în toată generalitatea, şi o mare parte a topologiei moderne este dedicată acestui lucru. Însă nu ar trebui să uităm niciodată faptul că, în marea majoritate a cazurilor care apar în studiul fenomenelor geometrice concrete, nu este indicat să se lucreze cu noţiuni a căror generalitate extremă poate crea dificultăţi nu neapărat necesare.

De fapt, teorema lui Jordan poate fi demonstrată cu ușurință pentru curbe destul de cumsecade, ca de pildă poligoane sau curbe cu tangente care se rotesc în mod continuu, și care intervin în cele mai importante probleme. Vom demonstra teorema pentru poligoane, în apendicele acestui capitol.

2. Problema celor patru culori

Din exemplul teoremei lui Jordan s-ar putea presupune că topologia se ocupă de furnizarea unor demonstrații riguroase pentru propoziții evidente, de care nici o persoană sănătoasă nu se poate îndoi. Dimpotrivă, există numeroase probleme topologice, unele chiar de formă foarte simplă, la care intuiția nu poate da vreun răspuns satisfăcător. Un exemplu de acest fel este renumita ,,problemă a celor patru culori".

Colorînd o hartă geografică, de obicei se atribuie culori diferite oricăror două țări care au frontieră comună. Experimental, s-a stabilit că orice hartă, oricîte țări ar conține și oricum ar fi situate, poate fi colorată doar cu *patru* culori diferite. Se poate ușor observa că un număr mai mic de culori nu va fi suficient în toate cazurile. Fig. 129 reprezintă o insulă în mijlocul mării, care nu poate fi colorată în mod potrivit cu mai puțin de patru culori, deoarece conține patru țări, fiecare avînd frontieră comună cu celelalte trei.

Faptul că pînă acum nu a fost găsită nici o hartă, a cărei colorare să necesite mai mult de patru culori, sugerează următoarea teoremă: *Pentru orice împărțire a planului în regiuni care nu se suprapun, putem nota întotdeauna regiunile cu unul dintre numerele* 1, 2, 3, 4, *astfel încît două regiuni adiacente să nu aibă același număr.* Prin regiuni ,,adiacente" înțelegem regiuni care au un segment de frontieră în comun; două regiuni care se întîlnesc într-un singur punct sau numai într-un număr finit de puncte (ca de pildă, statele Colorado și Arizona) nu vor fi numite adiacente, deoarece nu va rezulta nici o confuzie, dacă ele vor fi colorate cu aceeași culoare.

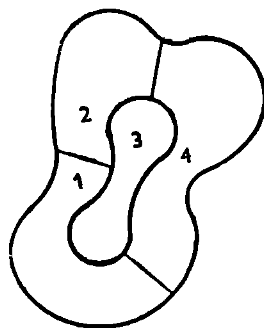


Fig. 129. Colorarea unei hărți.

Problema demonstrării acestor teoreme pare să fi fost propusă pentru prima dată de Moebius, în 1840, apoi de De Morgan, în 1850, și din nou, de Cayley, în 1878. O „demonstrație" a fost publicată de Kempe, în 1879, dar în 1890, Heawood a găsit o eroare în raționamentul lui Kempe. Verificînd demonstrația lui Kempe, Heawood a putut arăta că *cinci* culori sînt întotdeauna suficiente (o demonstrație a teoremei celor cinci culori este dată în apendicele acestui capitol). În ciuda eforturilor multor matematicieni celebri, situația este următoarea : s-a *demonstrat* că cinci culori sînt suficiente pentru orice hartă și se *presupune* că sînt suficiente patru culori. Dar, ca și în cazul celebrei teoreme a lui Fermat (cf. p. 57), nu s-a putut da nici o demonstrație acestei ipoteze și nici un exemplu care să o contrazică, astfel încît ea rămîne una dintre marile probleme nerezolvate ale matematicii. Teorema celor patru culori a fost demonstrată pentru toate hărțile care conțin mai puțin de treizeci și opt de regiuni. Datorită acestui fapt, pare că chiar dacă teorema generală este falsă, ea nu poate fi infirmată printr-un exemplu foarte simplu.

În problema celor patru culori, hărțile pot fi desenate fie în plan, fie pe suprafața unei sfere. Cele două cazuri sînt echivalente : Orice hartă de pe sferă poate fi reprezentată pe plan, efectuînd o mică gaură în interiorul uneia din regiunile *A*, deformînd apoi suprafața rămasă și așternînd-o pe un plan, ca în demonstrația teoremei lui Euler. Harta plană care rezultă va fi aceea a unei „insule", formată din regiunile rămase, înconjurată de o „mare", formată din regiunea A. Reciproc, inversînd acest proces, orice hartă din plan poate fi reprezentată pe sferă și de aceea ne putem restrînge la hărți de pe sferă. Mai mult, deoarece deformările regiunilor și ale frontierelor lor nu influențează problema, putem presupune că frontiera oricărei regiuni este un poligon închis și simplu, format din arce de cerc. Chiar astfel „regularizată", problema rămîne nerezolvată. Dificultățile care apar aici, spre deosebire de acelea care intervin în teorema lui Jordan, nu provin din generalitatea noțiunilor de regiune și curbă.

Un fapt remarcabil, legat de problema celor patru culori, este acela că pentru suprafețe mai complicate decît planul sau sfera, teoremele corespunzătoare au fost demonstrate. Ne apare astfel ca un lucru destul de paradoxal faptul că analiza unor suprafețe geometrice mai complicate pare a fi mai ușoară din acest punct de vedere decît aceea a unor suprafețe mai simple. De exemplu, pe suprafața unui tor (fig. 123), a cărui formă este aceea a unui covrig, sau a unei camere de automobil umflate, s-a arătat că orice hartă poate fi colorată cu ajutorul a șapte culori și s-au putut construi hărți care conțin șapte regiuni, fiecare dintre ele fiind în contact cu celelalte șase.

3. Noțiunea de dimensiune

Noțiunea de dimensiune nu prezintă dificultăți atît timp cît ne ocupăm numai de figuri geometrice simple, ca de pildă puncte, drepte, triunghiuri și poliedre. Un singur punct, sau orice mulțime *finită* de puncte, are dimensiunea zero, un segment de dreaptă este unidimen-

sional, în timp ce suprafața unui triunghi sau a unei sfere este bidimensională. Mulțimea punctelor dintr-un cub solid este tridimensională. Dar cînd încercăm să extindem această noțiune la mulțimi punctuale mai generale, apare necesitatea unei definiții precise. Ce dimensiune trebuie atribuită mulțimii R, formată din toate punctele de pe axa numerică, ale căror coordonate sînt numere *raționale*? Mulțimea punctelor raționale este densă pe dreaptă și de aceea ar putea fi considerată unidimensională, ca și dreapta însăși. Pe de altă parte, există goluri iraționale între orice pereche de puncte raționale, ca și între oricare două puncte ale unei mulțimi punctuale finite, astfel încît s-ar putea considera de asemenea că dimensiunea mulțimii R este egală cu zero.

O problemă mai complicată apare dacă încercăm să atribuim o dimensiune următoarei mulțimi ciudate, considerată pentru prima dată de Cantor. Din segmentul unitate extrageți treimea mijlocie, formată din toate punctele x, astfel încît $1/3 < x < 2/3$. Să notăm cu C_1 mulțimea rămasă. Acum, din C_1 să extragem treimea mijlocie a fiecăruia din cele două segmente și să notăm cu C_2 mulțimea rămasă. Repetînd acest proces de îndepărtare a treimii mijlocii a

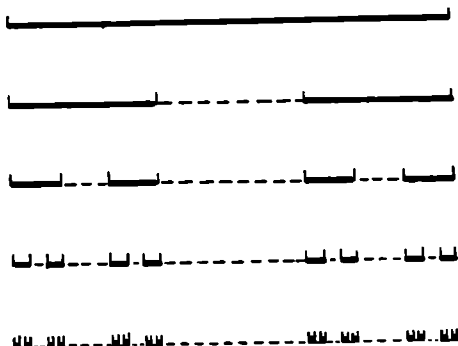


Fig. 130. Mulțimea punctuală a lui Cantor

fiecăruia dintre cele patru intervale ale lui C_2, obținem o mulțime C_3, și, continuînd în același mod, obținem mulțimile C_4, C_5, C_6, ... Să notăm cu C mulțimea punctelor de pe segmentul unitate, care rămîn după ce au fost îndepărtate toate aceste intervale, adică C este mulțimea punctelor comune șirului infinit de mulțimi C_1, C_2, ... Deoarece la primul pas am îndepărtat un interval de lungime 1/3, la al doilea pas două intervale, fiecare de lungime $1/3^2$ etc., lungimea totală a segmentelor extrase este egală cu

$$1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{3^2} + 2^2 \cdot \frac{1}{3^3} + \ldots = \frac{1}{3}\left(1 + \left(\frac{2}{3}\right) + \left(\frac{2}{3}\right)^2 + \ldots\right).$$

Seria infinită din paranteză este o serie geometrică, a cărei sumă este egală cu $1/(1 - 2/3) = 3$. Deci lungimea totală a segmentelor extrase este egală cu 1; și totuși mai rămîn puncte în mulțimea C. De exemplu, punctele 1/3, 2/3, 1/9, 2/9, 7/9, 8/9, ... prin care am efectuat

secţiunile succesive, aparţin mulţimii C. De fapt, este uşor de arătat că C constă din acele puncte x, a căror dezvoltare în fracţie triadiacă infinită poate fi scrisă sub forma

$$x = \frac{a_1}{3} + \frac{a_2}{3^2} + \frac{a_3}{3^3} + \ldots + \frac{a_i}{3^i} + \ldots,$$

unde fiecare a_i este egal fie cu 0, fie cu 2, în timp ce dezvoltarea triadică a fiecărui punct extras va avea cel puţin unul dintre numerele a_i egal cu 1.

Care trebuie să fie dimensiunea mulţimii C? Procedeul diagonal folosit la demonstrarea nenumerabilităţii mulţimii tuturor numerelor reale poate fi modificat, astfel încît să dea acelaşi rezultat şi pentru mulţimea C. Ar părea deci că mulţimea C trebuie să fie unidimensională. Şi totuşi C nu conţine nici un interval complet, oricît de mic ar fi el, aşa încît C ar putea fi considerată zero dimensională, ca şi o mulţime finită de puncte. În acelaşi mod ne-am putea întreba dacă mulţimea punctelor din *plan*, obţinută ridicînd un segment de lungime unitate, perpendicular pe axa $0X$, în fiecare punct raţional, sau în fiecare punct al mulţimii lui Cantor C, trebuie considerată a fi unidimensională sau bidimensională.

Poincaré (în 1912) a fost acela care a atras atenţia, pentru prima dată, asupra necesităţii unei analize mai profunde şi a unei definiţii precise a noţiunii de dimensiune. Poincaré a observat că dreapta este unidimensională, pentru că putem separa orice două puncte ale ei, secţionînd-o într-un singur punct (care este de dimensiune 0), în timp ce planul este bidimensional, deoarece pentru a separa o pereche de puncte din plan trebuie să trasăm o curbă închisă (de dimensiune 1). Acest fapt sugerează natura inductivă a dimensiunii: un spaţiu este n-dimensional, dacă două puncte oarecare ale lui pot fi separate prin extragerea unei submulţimi $(n - 1)$-dimensionale, şi dacă o submulţime de dimensiune mai mică nu va face acelaşi serviciu pentru orice pereche de puncte. O definiţie inductivă a dimensiunii mai este conţinută implicit în *Elementele* lui Euclid, în care o figură unidimensională este aceea a cărei frontieră constă din puncte, o figură bidimensională este aceea care are frontiera formată din curbe, iar o figură tridimensională este aceea care are frontiera formată din suprafeţe.

În ultimii ani a fost dezvoltată o teorie vastă a dimensiunii. O definiţie a dimensiunii începe prin definirea precisă a noţiunii de „mulţime punctuală de dimensiune 0" Orice mulţime *finită* de puncte are proprietatea că orice punct al mulţimii poate fi inclus într-o regiune a spaţiului, care poate fi făcută oricît de mică vrem şi care nu conţine nici un punct al mulţimii pe frontieră. Această proprietate se ia ca definiţie a dimensiunii nule. Pentru comoditate, spunem că mulţimea vidă, care nu conţine nici un punct, are dimensiunea egală cu -1. Atunci, o mulţime punctuală S este de dimensiune 0, dacă ea nu este de dimensiune -1 (adică S conţine cel puţin un punct), şi dacă orice punct al lui S poate fi inclus într-o regiune oricît de mică, a cărei frontieră intersectează pe S după o mulţime de dimensiune -1 (adică a cărei frontieră nu conţine puncte ale lui S). De exemplu, mulţimea punctelor raţionale de pe dreaptă este de dimensiune 0, deoarece orice punct raţional este centrul unui interval oricît de mic, cu extremităţi iraţionale. Mulţimea lui Cantor C are şi ea dimensiunea 0, deoarece, ca şi mulţimea punctelor raţionale, se obţine prin extragerea unei mulţimi dense de puncte de pe dreaptă.

Pînă acum am definit numai noţiunile de dimensiune -1 şi de dimensiune 0. Definiţia dimensiunii 1 se impune imediat: o mulţime S de puncte este de dimensiune 1, dacă ea nu este de dimensiune -1 sau 0 şi dacă fiecare punct al lui S poate fi inclus în interiorul unei regiuni

oricît de mici, a cărei frontieră intersectează pe S după o mulţime de dimensiune 0. Un segment de dreaptă are această proprietate, deoarece frontiera oricărui interval este o pereche de puncte care este de dimensiune 0, conform definiţiei precedente. Mai mult, continuînd în acelaşi mod, putem defini succesiv noţiunile de dimensiune 2, 3, 4, 5, ..., fiecare bazîndu-se pe noţiunile precedente. Astfel, o mulţime S va fi de dimensiune n, dacă ea nu este de dimensiune mai mică şi dacă orice punct al lui S poate fi inclus în interiorul unei regiuni oricît de mici, a cărei frontieră intersectează pe S după o mulţime de dimensiune $n-1$. De exemplu, planul este de dimensiune 2, deoarece orice punct al planului poate fi incus în interiorul unui cerc oricît de mic, a cărui circumferinţă are dimensiunea 1*. Nici o mulţime de puncte din spaţiul obişnuit nu poate avea o dimensiune mai mare decît 3, deoarece orice punct al spaţiului este centrul unei sfere oricît de mici, a cărei suprafaţă are dimensiunea 2. Însă în matematica modernă, cuvîntul „spaţiu" este folosit pentru desemnarea oricărui sistem de obiecte pentru care este definită o noţiune de „distanţă" sau „vecinătate" (cf. p. 334), iar aceste „spaţii" abstracte pot avea dimensiuni mai mari decît 3. Un exemplu simplu este *spaţiul cartezian n-dimensional*, ale cărui „puncte" sînt sisteme ordonate de n numere reale:

$$P = (x_1, x_2, x_3, \ldots, x_n),$$

$$Q = (y_1, y_2, y_3, \ldots, y_n),$$

„distanţa" dintre punctele P şi Q fiind definită prin

$$d(P, Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}.$$

Se poate arăta că acest spaţiu are dimensiunea n. Un spaţiu care nu are dimensiunea n, oricare ar fi întregul n, este prin definiţie de dimensiune infinită. Se cunosc multe exemple de astfel de spaţii.

Unul dintre cele mai interesante fapte din teoria dimensiunii este următoarea proprietate caracteristică a figurilor bi-, tri- sau, în general, n-dimensionale. Să considerăm mai întîi cazul bidimensional. Dacă o figură bidimensională simplă este subdivizată în regiuni destul de mici (fiecare fiind considerată împreună cu frontiera ei), atunci vor exista în mod necesar puncte care aparţin *la trei sau mai multe* regiuni, *oricare ar fi forma lor*. În plus, *vor exista subdiviziuni* ale figurii, în care fiecare punct aparţine la *cel mult* trei regiuni ale subdiviziunii. Astfel, dacă figura bidimensională este un pătrat, ca în fig. 131, atunci un punct va aparţine celor trei regiuni 1, 2 şi 3, în timp ce pentru această subdiviziune particulară nici un punct nu aparţine la mai mult de trei regiuni. În mod similar, se poate arăta că, în cazul tridimensional, dacă acoperim un volum cu volume destul de mici, vor exista întotdeauna puncte comune la cel puţin patru dintre acestea, în timp ce pentru o subdiviziune aleasă în mod potrivit, nu mai mult de patru regiuni vor avea un punct comun.

* Aceasta nu pretinde a fi o demonstraţie riguroasă a faptului că planul este de dimensiune 2 în conformitate cu definiţia dată, deoarece ea presupune că circumferinţa unui cerc este de dimensiune 1, şi se presupune cunoscut faptul că planul nu este de dimensiune 0 sau 1. Însă se poate da o demonstraţie a acestor fapte şi a analogelor lor în dimensiuni mai mari. Această demonstraţie arată că definiţia dimensiunii unei mulţimi punctuale în cazul general nu o contrazice pe cea uzuală, pentru mulţimile simple.

Aceste observații sugerează următoarea teoremă, datorată lui Lebesgue și Brouwer: Dacă o figură n-dimensională este acoperită într-un mod oarecare cu subregiuni suficient de mici, atunci vor exista puncte care aparțin la cel puțin $n + 1$ subregiuni; mai mult, se poate găsi întotdeauna o acoperire cu regiuni oricît de mici, astfel încît nici un punct să nu aparțină la mai mult de $n + 1$ regiuni. Datorită metodei de acoperire considerată aici, aceasta este cu-

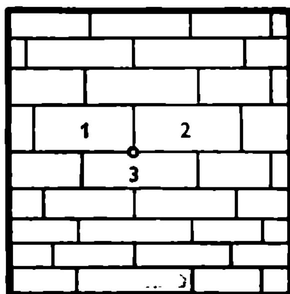


Fig. 131. Teorema pavajului

noscută sub numele de teorema „pavajului". Ea caracterizează dimensiunea oricărei figuri geometrice: figurile pentru care ea este valabilă sînt n-dimensionale, în timp ce toate celelalte au o altă dimensiune. Din acest motiv, ea poate fi luată ca *definiție* a dimensiunii, așa cum fac unii autori.

Dimensiunea oricărei mulțimi este o caracteristică topologică a mulțimii; două figuri de dimensiuni diferite nu pot fi topologic echivalente. Aceasta este celebra teoremă topologică a „invarianței dimensiunii", care devine mai importantă dacă facem comparația cu faptul enunțat la p.101, în virtutea căruia mulțimea punctelor unui pătrat are același număr cardinal ca și mulțimea punctelor de pe un segment de dreaptă. Corespondența care a fost definită acolo nu este topologică, deoarece condițiile de continuitate nu sînt respectate.

4. O teoremă de punct fix

În aplicațiile topologiei la alte ramuri ale matematicii, teoremele de „punct fix" joacă un rol important. Un exemplu tipic este următoarea teoremă a lui Brouwer. Ea este mai puțin evidentă intuitiv decît alte teoreme topologice.

Să considerăm un disc circular în plan. Prin aceasta înțelegem regiunea formată din interiorul unui cerc, împreună cu circumferința sa. Să presupunem că punctele acestui disc sînt supuse unei transformări continue (care nu trebuie să fie neapărat biunivocă), prin care fiecare punct rămîne în interiorul cercului, chiar dacă își schimbă poziția. De exemplu, un disc subțire din cauciuc poate fi strîns, întors, împăturit, întins sau deformat în orice mod, dar astfel încît poziția finală a fiecărui punct al discului să se găsească în interio-

rul circumferinței inițiale. Un alt exemplu se obține dacă presupunem că lichidul conținut într-un pahar este pus în mișcare prin agitare, astfel încît particulele care se află la suprafață să rămînă la suprafață : atunci, în orice moment, poziția particulelor de pe suprafață definește o transformare continuă a distribuției inițiale a particulelor. Teorema lui Brouwer se enunță acum în felul următor : *Orice transformare de acest fel are cel puțin un punct fix* : adică există cel puțin un punct, a cărui poziție după transformare este aceeași cu poziția inițială. (În exemplul cu suprafața lichidului, punctul fix în general se va schimba cu timpul, cu toate că în cazul unor rotații circulare simple, centrul rămîne mereu fixat.) Demonstrația existenței unui punct fix este un exemplu caracteristic de raționament folosit pentru stabilirea multor teoreme topologice.

Să considerăm discul înainte și după transformare și să presupunem, contrar enunțului teoremei, că *nici un* punct nu rămîne fix, astfel încît, prin transformare, fiecare punct se deplasează în alt punct din interior sau de pe cerc. Oricărui punct P din discul inițial să-i asociem o săgeată sau un ,,vector'', care indică direcția PP', unde P' este imaginea lui P prin transformare. În fiecare punct al discului avem o astfel de săgeată, deoarece fiecare punct a fost presupus deplasat în altă poziție. Să considerăm acum punctele de pe frontiera cercului, împreună cu vectorii asociați. Toți acești vectori sînt îndreptați spre interiorul cercului deoarece, prin ipoteză, nici un punct nu este transformat într-un punct exterior cercului. Să începem cu un

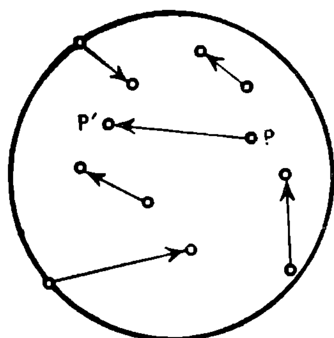


Fig. 132. Vectorii de transformare

punct oarecare P_1 de pe frontieră și să ne deplasăm de-a lungul cercului în sens opus acelor unui ceasornic. Făcînd acest lucru, direcția vectorului se va modifica, deoarece vectorii asociați punctelor de pe frontieră sînt variabili. Direcțiile acestor vectori pot fi indicate desenînd săgeți paralele, care pornesc dintr-un singur punct al planului. Observăm că colind cercul, pornind din P_1 și revenind în P_1, vectorul se rotește și revine în

poziția inițială. Să numim „indice" al vectorilor de pe cerc numărul de rotații complete efectuate de acest vector; mai precis, definim indicele ca fiind *suma algebrică* a diferitelor modificări ale unghiurilor vectorilor, astfel încît orice rotație efectuată în sensul acelor de ceasornic este luată cu semnul minus, în timp ce orice rotație efectuată în sensul opus acelor de ceasornic este considerată pozitivă. Indicele este rezultatul net, care *a priori* poate fi unul din numerele

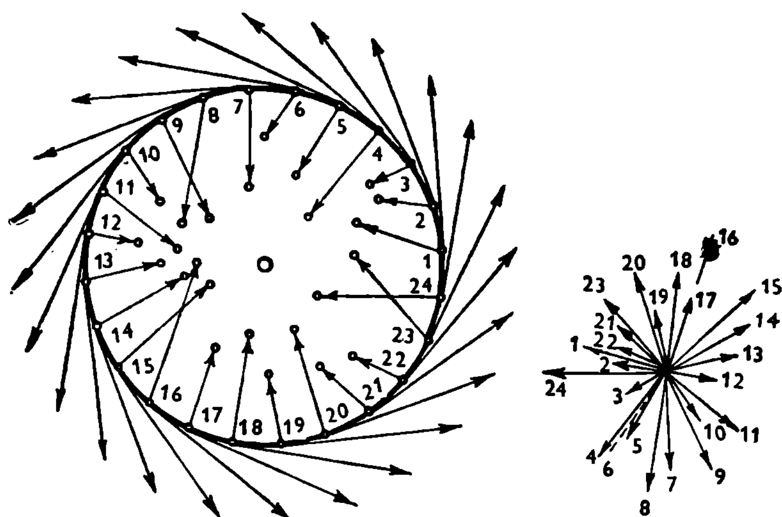


Fig. 133.

0, $\pm$ 1, $\pm$ 2, $\pm$ 3, ... corespunzător unei variații totale a unghiului, egală cu 0, $\pm$ 360°, $\pm$ 720°, ... Afirmăm acum că *indicele este egal cu* 1, adică variația totală a direcției săgeții este egală exact cu o rotație pozitivă. Pentru a arăta acest lucru, reamintim că vectorul de transformare al oricărui punct P de pe cerc este îndreptat întotdeauna spre interiorul cercului și niciodată în lungul tangentei. Dar dacă acest vector de transformare se rotește cu un unghi total, diferit de unghiul total cu care se rotește vectorul *tangent*, (care este de 360°, pentru că vectorul tangent efectuează evident o rotație pozitivă completă), atunci diferența dintre unghiurile totale cu care vectorul tangent și vectorul de transformare se rotesc va fi un multiplu nenul de 360°, deoarece fiecare efectuează un număr întreg de rotații complete. Deci vectorul de transformare trebuie să se rotească complet în jurul tangentei cel puțin o dată, în timpul circuitului complet care pornește din P_1 și se termină în P_1.

Dar deoarece vectorul de transformare şi cel tangent se rotesc continuu, într-un anume punct al circumferinţei vectorul de transformare trebuie să fie îndreptat în lungul celui tangent, ceea ce, după cum am văzut, este imposibil.

Dacă considerăm acum un cerc concentric cu circumferinţa discului, conţinut în interior, împreună cu vectorii de transformare corespunzători de pe acest cerc, atunci indicele vectorilor de transformare de pe cerc trebuie să fie egal tot cu 1. Acest rezultat se justifică prin faptul că atunci cînd trecem în mod continuu de pe circumferinţă pe alt cerc concentric, indicele trebuie să varieze continuu, deoarece direcţiile vectorilor de transformare variază continuu în interiorul discului. Dar indicele poate lua doar valori întregi şi de aceea trebuie să fie mereu egal cu valoarea iniţială 1, deoarece un salt de la 1 la altă valoare întreagă ar fi o discontinuitate în comportarea indicelui. (Concluzia că o mărime care variază continuu, dar care poate lua doar valori întregi, este în mod necesar o constantă, este un raţionament tipic matematic, care intervine în multe demonstraţii.) Astfel, putem găsi un cerc concentric, oricît de mic, pentru care indicele corespunzător al vectorilor de transformare este egal cu 1. Însă acest lucru este imposibil, deoarece datorită continuităţii vectorilor de transformare, într-un cerc suficient de mic, toţi vectorii vor indica aproximativ aceeaşi direcţie ca şi vectorul din centrul cercului. Astfel, variaţia totală a unghiurilor lor poate fi făcută oricît de mică vrem, de pildă mai mică decît 10°, luînd un cerc suficient de mic. Prin urmare, indicele care trebuie să fie un întreg, va fi nul. Această contradicţie arată că ipoteza noastră iniţială, după care nu există puncte fixe ale transformării, nu este acceptabilă şi aceasta încheie demonstraţia.

Teorema pe care tocmai am demonstrat-o îşi păstrează valabilitatea nu numai pentru disc, ci şi pentru o regiune triunghiulară, sau pătrată, sau orice altă suprafaţă, care este imaginea unui disc printr-o transformare topologică. Într-adevăr, dacă A este orice figură legată de un disc printr-o transformare biunivocă şi bicontinuă, atunci o transformare continuă a lui A în ea însăşi, care nu ar avea puncte fixe, ar defini o transformare continuă a discului în el însuşi, fără puncte fixe, ceea ce, după cum am văzut, este imposibil. Teorema îşi păstrează valabilitatea în spaţiul tridimensional, pentru sfere sau cuburi solide, însă demonstraţia nu mai este atît de simplă.

Cu toate că teorema de punct fix a lui Brouwer pentru disc nu este atît de evidentă din punct de vedere intuitiv, este uşor de arătat că ea este o consecinţă imediată a următorului fapt, al cărui adevăr este intuitiv evident: *este imposibil să transformăm în mod continuu un disc circular în circumferinţa sa, astfel încît fiecare punct al circumferinţei să rămînă fix*. Vom arăta că existenţa unei transformări fără puncte fixe, a discului în el însuşi, ar contrazice acest fapt. Să presupunem că $P \to P'$ ar fi o astfel de transformare; pentru orice punct P al discului am putea duce o săgeată care porneşte din P', trece prin P şi se termină pe circumferinţă într-un punct P^*. Atunci, transformarea $P \to P^*$ ar fi o transformare continuă a întregului disc în circumferinţa sa şi ar lăsa fix orice punct al circumferinţei, contrar ipotezei că o astfel de

transformare este imposibilă. Un raţionament similar poate fi folosit pentru stabilirea teoremei lui Brouwer în trei dimensiuni, pentru sfera sau cubul solid.

Este uşor de văzut că unele figuri geometrice admit transformări continue în ele însele, fără puncte fixe. De exemplu, regiunea inelară, cuprinsă între două cercuri concentrice, admite ca transformare continuă fără puncte fixe orice rotaţie în jurul centrului, cu un unghi care nu este multiplu de 360°. Suprafaţa unei sfere admite ca transformare continuă fără puncte fixe pe aceea care constă din transformarea fiecărui punct în punctul diametral opus. Însă se poate arăta, printr-un raţionament asemănător cu acela pe care l-am folosit pentru disc, că orice transformare continuă, care nu transformă nici un punct în punctul diametral opus (de pildă orice deformare mică) are un punct fix.

Teoreme de punct fix, ca acelea de mai sus, furnizează o metodă puternică pentru demonstrarea multor „teoreme de existenţă", matematice care la prima vedere nu par a avea un caracter geometric. Un exemplu celebru este o teoremă de punct fix, conjecturată de Poincaré în 1912, cu puţin înainte de moartea sa. Această teoremă are drept consecinţă imediată existenţa unui număr infinit de orbite periodice, în problema restrînsă a celor trei corpuri. Poincaré nu a putut confirma ipoteza sa şi de aceea a fost o realizare majoră a matematicii americane faptul că, în anul următor, G. D. Birkhoff a reuşit să dea o demonstraţie. De atunci, metodele topologice au fost aplicate cu mult succes în studiul comportării calitative a sistemelor dinamice.

5. Noduri

Ca ultim exemplu, se poate scoate în evidenţă faptul că studiul nodurilor prezintă probleme matematice dificile, cu caracter topologic. Un nod se formează încolăcind o sfoară şi unind apoi capetele ei. Curba închisă care se obţine este o figură geometrică care rămîne esenţial neschimbată dacă o defor-

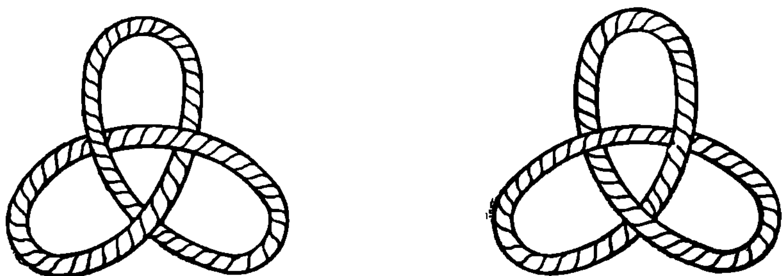


Fig. 134. Noduri topologice echivalente, care nu sînt deformabile unul într-altul

măm, trăgînd-o sau îndoind-o fără a rupe sfoara. Cum putem da însă o caracterizare intrinsecă, cu ajutorul căreia să putem deosebi o curbă închisă înnodată din spaţiu, de o curbă neînnodată, cum ar fi cercul? Răspunsul nu este cîtuşi de puţin simplu, şi cu atît mai puţin simplă este analiza matematică

completă a diferitelor feluri de noduri și a deosebirilor dintre ele. Chiar și pentru cel mai simplu caz, s-a dovedit că acesta este un lucru dificil. Să considerăm cele două noduri în formă de trifoi, arătate în fig. 134. Aceste două noduri sînt „imagini în oglindă", complet simetrice unul față de altul, și sînt topologic echivalente, fără a fi identice. Se pune problema dacă este posibil să deformăm unul din aceste noduri în celălalt, într-un mod continuu. Răspunsul este negativ, însă demonstrarea acestui fapt necesită o mult mai bună cunoaștere a tehnicii topologiei și teoriei grupurilor decît cea care poate fi prezentată aici.

§ 4. CLASIFICAREA TOPOLOGICĂ A SUPRAFEȚELOR

1. Genul unei suprafețe

Multe fapte topologice simple, dar importante, apar în studiul suprafețelor bidimensionale. De exemplu, să comparăm suprafața unei sfere cu aceea a unui tor. Din fig. 135 se vede că cele două suprafețe se deosebesc fundamental: pe sferă, ca și în plan, orice curbă simplă închisă, ca de pildă C, separă suprafața în două părți, însă pe tor există curbe închise, ca de pildă C', care nu separă suprafața în două părți. A spune că C separă sfera în două părți, înseamnă că, dacă sfera este tăiată în lungul lui C, ea se va descompune în două bucăți deosebite și nelegate sau, ceea ce este același lucru, putem găsi două puncte pe sferă, astfel încît orice curbă de pe sferă, care le unește, trebuie să inter-

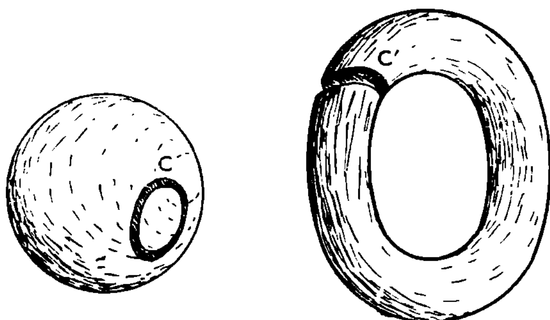


Fig. 135. Tăieturi pe sferă și tor

secteze pe C. Pe de altă parte, dacă torul este tăiat în lungul curbei închise C', suprafața obținută rămîne dintr-o bucată: orice punct de pe suprafață poate fi unit cu oricare alt punct printr-o curbă care nu intersectează pe C'. Această deosebire dintre sferă și tor face ca cele două suprafețe să fie topologic

distincte şi arată că este imposibil să deformăm în mod continuu pe una în cealaltă.

Să considerăm acum suprafaţa cu două găuri, arătată în fig. 136. Pe această suprafaţă putem trasa *două* curbe închise care nu se intersectează, A şi B, şi care nu separă suprafaţa. Torul este separat întotdeauna în două părţi prin două curbe de acest fel. Pe de altă parte, *trei* curbe închise care nu se intersectează separă întotdeauna suprafaţa cu două găuri.

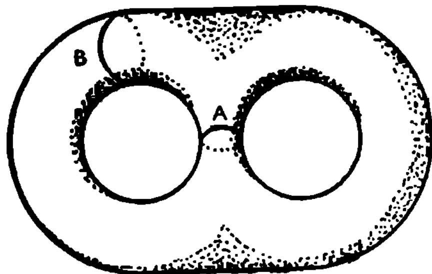


Fig. 136. O suprafaţă de genul doi

Aceste fapte sugerează definirea *genului* unei suprafeţe, ca fiind cel mai mare număr de curbe simple închise, care nu se intersectează şi care pot fi trasate pe suprafaţă, fără a o separa. Genul sferei este 0, acela al torului este 1, în timp ce acela al suprafeţei din fig. 136 este 2. O suprafaţă asemănătoare, cu p găuri, are genul p. Genul este o proprietate topologică a suprafeţei şi rămîne neschimbat dacă suprafaţa este deformată. Reciproc, se poate arăta (omitem aici demonstraţia) că dacă două suprafeţe închise au acelaşi gen, atunci una poate fi deformată în cealaltă, astfel încît genul $p = 0, 1, 2, \ldots$ al unei suprafeţe închise o caracterizează complet din punct de vedere topologic. (Pre-

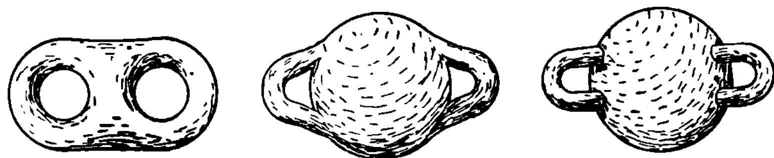


Fig. 137. Suprafeţe de genul doi

supunem că suprafeţele considerate sînt suprafeţe închise „bilatere" obişnuite. În secţiunea 3 a acestui paragraf vom considera suprafeţe „unilatere".) De exemplu, covrigul cu două găuri şi sfera cu două „toarte" din fig. 137 sînt ambele suprafeţe închise de genul 2 şi este limpede că fiecare dintre aceste

suprafețe poate fi deformată continuu în cealaltă. Deoarece covrigul cu p găuri, sau echivalentul său, sfera cu p toarte, sînt de genul p, putem lua fiecare dintre aceste suprafețe ca reprezentant topologic al tuturor suprafețelor închise de genul p.

2. Caracteristica lui Euler a unei suprafețe

Să presupunem că o suprafață închisă S de genul p este împărțită într-un număr de regiuni prin alegerea unui număr de vîrfuri pe S, unite prin arce de curbe. Vom arăta că

$$V - M + F = 2 - 2p, \tag{1}$$

unde V este numărul vîrfurilor, M numărul arcelor, iar F numărul regiunilor. Numărul $2 - 2p$ se numește *caracteristica lui Euler* a suprafeței. Am văzut deja că pentru sferă avem $V - M + F = 2$, ceea ce concordă cu (1), deoarece sfera este de gen zero.

Pentru a demonstra formula generală (1), putem presupune că S este o sferă cu p toarte, deoarece așa cum am spus, orice suprafață de genul p poate fi deformată continuu într-o suprafață de acest fel și în timpul acestei deformări numerele $V - M + F$ și $2 - 2p$ nu se modifică. Vom alege deformarea astfel, încît curbele închise A_1, A_2, B_1, B_2, ... în lungul cărora toartele se îmbină cu sfera, să fie formate din arce ale împărțirii efectuate. (Ne referim la fig. 138 care ilustrează demonstrația pentru cazul $p = 2$.)

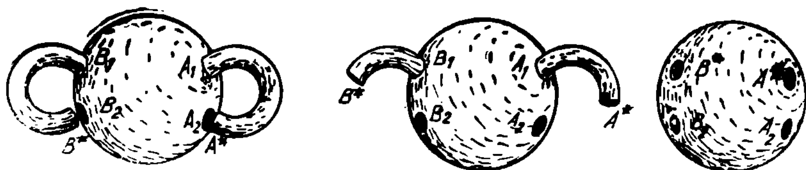


Fig. 138.

Să tăiem acum suprafața S în lungul curbelor A_2, B_2, ... și să îndreptăm toartele. Fiecare toartă va avea o muchie liberă, mărginită de o nouă curbă A^*, B^*, ..., cu același număr de vîrfuri și de arce ca și A_2, B_2, ... Deci, $V - M + F$ nu se va modifica, deoarece vîrfurile suplimentare compensează arcele suplimentare și nu se formează noi regiuni. Apoi deformăm suprafața, netezind toartele proeminente, pînă ce obținem o suprafață care este o sferă din care au fost extrase $2p$ regiuni. Deoarece $V - M + F$ este egal cu 2 pentru orice împărțire a sferei întregi, avem:

$$V - M + F = 2 - 2p$$

pentru sfera cu $2p$ regiuni extrase, și deci și pentru sfera inițială cu p toarte, așa cum trebuia să demonstrăm.

Fig. 121 ilustrează aplicarea formulei (1) unei suprafețe S formată din poligoane plane. Această suprafață poate fi deformată continuu într-un tor, astfel încît genul p este egal cu 1 și $2 - 2p = 2 - 2 = 0$. Conform formulei (1), avem

$$V - M + F = 16 - 32 + 16 = 0.$$

Exercițiu: Subdivizați covrigul cu două găuri din fig. 137 în regiuni, și arătați că $V - M + F = -2$.

3. Suprafețe unilatere

O suprafață obișnuită are două fețe. Acest lucru este adevărat atît pentru suprafețe închise, ca de pildă sfera sau torul, ca și pentru suprafețe mărginite de curbe, cum ar fi discul sau torul, din care a fost extrasă o bucată. Cele două fețe ale unei astfel de suprafețe pot fi vopsite cu culori diferite, pentru a le deosebi una de alta. Dacă suprafața este închisă, cele două culori nu se întîlnesc niciodată. Dacă suprafața este mărginită de curbe, cele două culori se întîlnesc doar în lungul lor. O ploșniță care s-ar tîrî în lungul unei astfel de suprafețe, fără să traverseze eventualele curbe frontieră, ar rămîne mereu pe aceeași față.

Moebius a făcut descoperirea surprinzătoare că există suprafețe cu *o singură* față. Cea mai simplă suprafață de acest fel este așa-numita bandă a lui Moebius, formată dintr-o fîșie dreptunghiulară lungă de hîrtie, răsucită și lipită la capete, așa cum se arată în fig. 139. O ploșniță care s-ar tîrî în lungul acestei suprafețe rămînînd mereu în mijlocul benzii, ar reveni în poziția inițială pe partea cealaltă. Banda lui Moebius are o singură margine, deoarece frontiera ei constă dintr-o singură curbă închisă. Suprafața bilateră obișnuită, obținută prin lipirea celor două capete ale unui dreptunghi, fără a-l răsuci în prealabil, are două curbe frontieră distincte. Dacă această bandă este tăiată în lungul curbei centrale, se obțin două benzi diferite, de același fel. Însă dacă tăiem banda lui Moebius în lungul aceleiași curbe (arătată în fig. 139), găsim că ea rămîne dintr-o bucată. Rareori se întîmplă ca o persoană nefamiliarizată cu banda lui Moebius să poată prezice această comportare, atît de contrară intuiției. Dacă suprafața care rezultă prin tăierea benzii lui Moebius în lungul curbei centrale este tăiată din nou în lungul mijlocului său, atunci obținem două benzi separate, dar înlănțuite.

Este un lucru fascinant jocul cu astfel de benzi, tăiate în lungul unor curbe paralele cu frontiera și aflate la o distanță de o doime, o treime etc. față de frontieră.

Frontiera unei benzi a lui Moebius este o curbă închisă, neînnodată, care poate fi deformată într-una plană, ca de pildă un cerc. În timpul deformării se poate permite ca banda să se intersecteze cu ea însăși, astfel încît se obține o

suprafață unilateră, care se intersectează pe ea însăși, arătată în fig. 140 și cunoscută sub numele de pălăria intersectată[1]. Locul după care ea se auto-intersectează este considerat ca fiind format din două curbe deosebite, fiecare

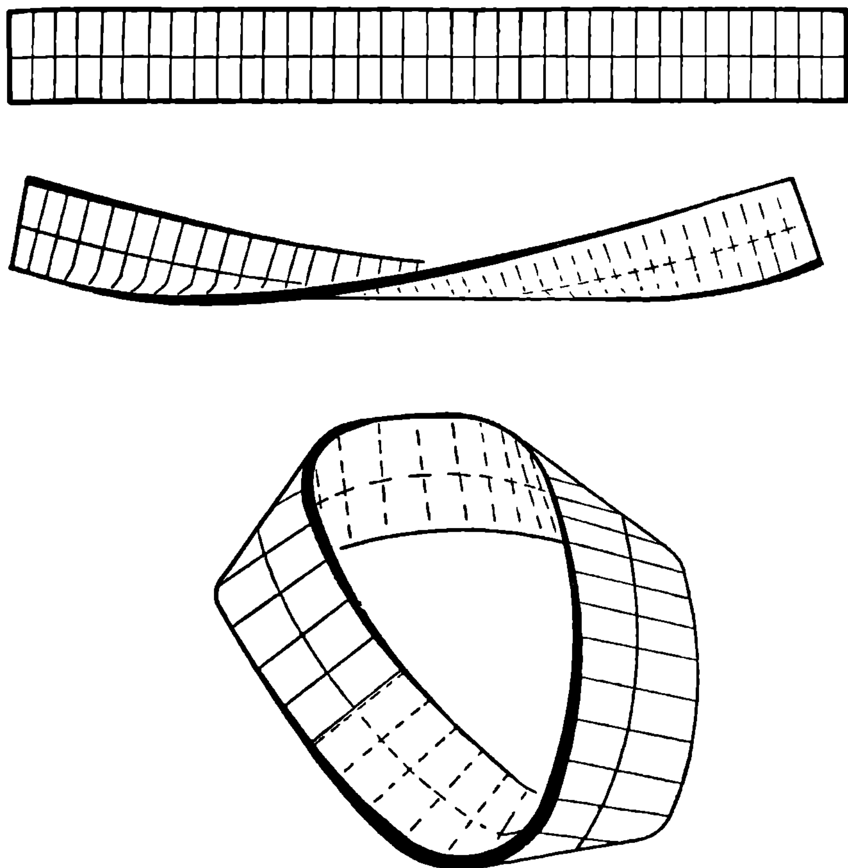


Fig. 139. Confecționarea unei benzi a lui Moebius

aparținînd uneia dintre cele două porțiuni ale suprafeței care se intersectează în acel loc. Unilateralitatea benzii lui Moebius se conservă, deoarece această proprietate este topologică ; o suprafață unilateră nu poate fi deformată în mod continuu într-o suprafață bilateră. Este destul de surprinzător faptul că defor-

[1] Termenul, neîncetățenit în limba română, este traducerea termenului *cross-cap*, folosit de autori. — *N.T.*

marea poate fi făcută în aşa fel, încît frontiera benzii lui Moebus să devină plană, de pildă triunghiulară, şi totuşi banda să nu se autointersecteze. Fig. 141 indică un astfel de model, datorat dr. B. Tuckermann; frontiera este un triunghi care formează o jumătate a unui pătrat diagonal al unui octaedru regulat; banda însăşi constă din şase feţe ale octaedrului şi patru triunghiuri dreptunghice, fiecare fiind un sfert dintr-nuplan diagonal.

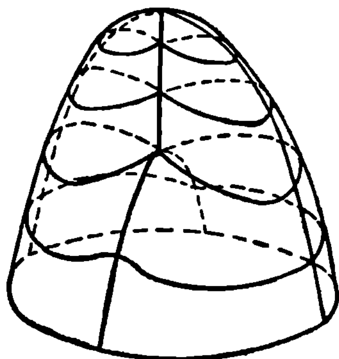


Fig. 140. Pălăria intersectată (cross-cap)

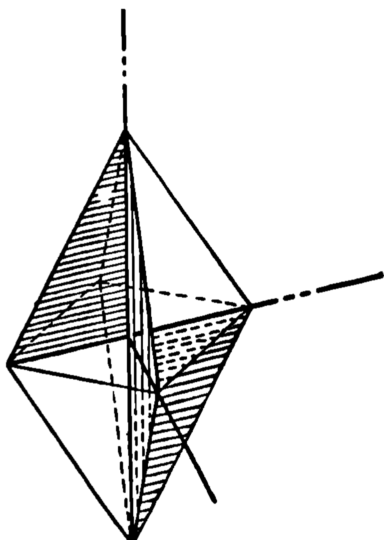


Fig. 141. Banda lui Moebius cu frontieră plană

O altă suprafaţă unilateră este „sticla lui Klein". Această suprafaţă este închisă, însă nu are nici interior, nici exterior. Ea este topologic echivalentă cu o pereche de pălării intersectate, ale căror frontiere coincid.

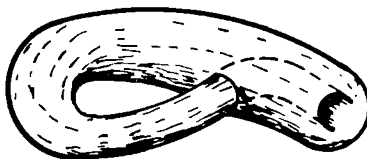


Fig. 142. Sticla lui Klein

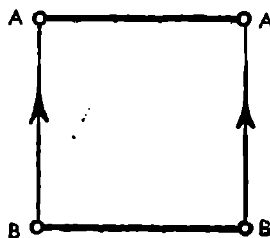
Se poate arăta că orice suprafaţă închisă *unilateră*, de genul $p = 1, 2, \ldots,$ este topologic echivalentă cu o sferă din care au fost extrase p discuri, care au fost înlocuite cu pălării intersectate. De aici rezultă, cu uşurinţă, că caracte-

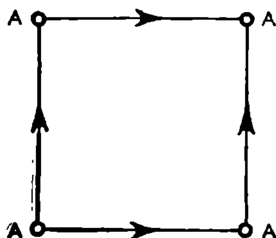
ristica lui Euler $V - M + F$ a unei astfel de suprafețe este legată de p prin egalitatea

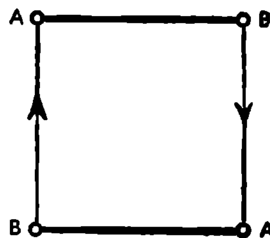
$$V - M + F = 2 - p.$$

Demonstrația este analogă cu cea dată pentru suprafețele bilatere. Mai întîi să arătăm că caracteristica lui Euler a unei pălării intersectate sau a benzii lui Moebius, este egală cu 0. Pentru a face acest lucru, observăm că secționînd o bandă a lui Moebius, care a fost sub divizată într-un număr de regiuni, obținem un dreptunghi, care conține în plus două vîrfuri, o muchie și are același număr de regiuni ca și banda lui Moebius. Pentru dreptunghi $V - M + F = 1$, după cum am demonstrat la p. 256. Deci pentru banda lui Moebius, $V - M + F = 0$. Ca exercițiu, cititorul poate completa demonstrația.

Este mult mai simplu să studiem natura topologică a suprafețelor de acest fel cu ajutorul poligoanelor plane, în care anumite perechi de laturi sînt iden-







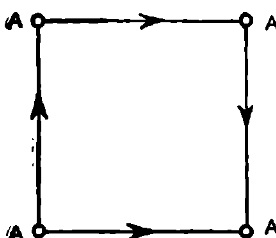


Fig. 143. Suprafețe închise definite prin punerea în corespondență a muchiilor unei figuri plane

tificate mintal (cf. cap. IV, apendice, secțiunea 3). În diagramele din fig. 143, săgețile paralele trebuie suprapuse — efectiv sau mintal — astfel încît să coincidă ca poziție și sens.

Această metodă de identificare poate fi folosită și pentru a defini și varietăți închise tridimensionale, analoge suprafețelor închise bidimensionale.

De exemplu, dacă identificăm punctele corespunzătoare ale fețelor opuse ale unui cub (fig. 144), obținem o varietate tridimensională închisă, numită torul tridimensional. Această varietate este topologic echivalentă cu spațiul cuprins

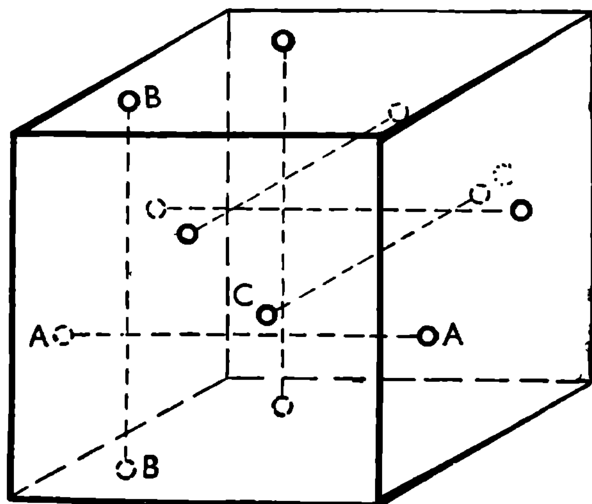


Fig. 144. Torul tridimensional definit prin identificări ale frontierei

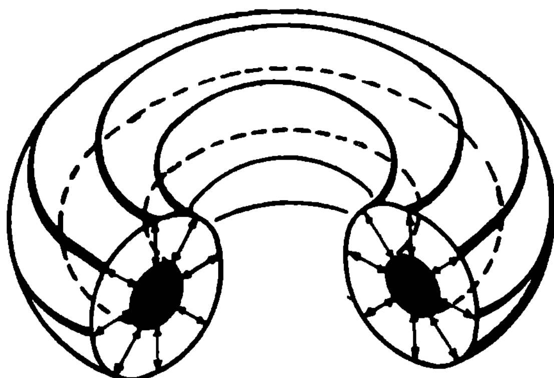


Fig. 145. O altă reprezentare a torului tridimensional (figura este secționată pentru a arăta identificările)

între două suprafețe toroidale concentrice, una în interiorul celeilalte, în care punctele corespunzătoare ale celor două suprafețe toroidale sînt identificate (fig. 145). Într-adevăr, ultima varietate se obține din cub, dacă două perechi de fețe identificate mintal sînt suprapuse.

APENDICE

*1. Teorema celor cinci culori

Pe baza formulei lui Euler, putem arăta că orice hartă de pe sferă poate fi colorată în mod potrivit, folosind cel mult cinci culori diferite. (Potrivit celor spuse la p. 263, se consideră că o hartă este colorată în mod potrivit, dacă două regiuni care au în comun o porțiune a frontierei lor sînt colorate cu culori diferite.) Ne vom limita la hărți ale căror regiuni sînt mărginite de poligoane simple, închise, formate din arce de cerc. Putem presupune, de asemenea, că în fiecare vîrf se întîlnesc trei arce; o astfel de hartă se va numi *regulată*. Putem face această presupunere deoarece, dacă înlocuim fiecare vîrf în care se întîlnesc mai mult de trei arce, printr-un cerc mic și dacă unim interiorul fiecărui cerc de acest fel cu una dintre regiunile care se întîlnesc în acest vîrf, obținem o nouă hartă, în care vîrfurile multiple sînt înlocuite printr-un număr de vîrfuri, fiecare fiind comun la trei arce. Noua hartă va conține același număr de regiuni ca și cea inițială. Dacă noua hartă, care este regulată, poate fi colorată în mod potrivit cu ajutorul a cinci culori, atunci contractînd cercurile și reducîndu-le la centrul lor, vom avea colorarea dorită a hărții inițiale. Astfel, este suficient să demonstrăm că orice hartă regulată de pe sferă poate fi colorată cu cinci culori.

Mai întîi arătăm că orice hartă regulată trebuie să conțină cel puțin un poligon cu mai puțin de șase laturi. Să notăm cu F_n numărul regiunilor cu n laturi dintr-o hartă regulată; atunci, dacă F este numărul total de regiuni, avem

$$F = F_2 + F_3 + F_4 + \dots . \tag{1}$$

Orice arc are două capete și în fiecare vîrf se întîlnesc trei arce. Prin urmare, dacă M este numărul de arce ale hărții, iar V este numărul de vîrfuri, atunci

$$2M = 3V. \tag{2}$$

Mai mult, o regiune mărginită de n arce are n vîrfuri și fiecare vîrf aparține la trei regiuni, astfel încît

$$2M = 3V = 2F_2 + 3F_3 + 4F_4 + \dots . \tag{3}$$

Din formula lui Euler avem

$$V - M + F = 2 \text{ sau } 6V - 6M + 6F = 12.$$

Din (2) vedem că $6V = 4M$, astfel încît $6F - 2M = 12$.
Deci din (1) și (3) obținem:

$$6(F_2 + F_3 + F_4 + \dots) - (2F_2 + 3F_3 + 4F_4 + \dots) = 12$$

sau

$$(6-2)F_2+(6-3)F_3+(6-4)F_4+(6-5)F_5+(6-6)F_6+$$
$$+(6-7)F_7+\ldots=12.$$

Deci cel puțin unul dintre termenii aflați în stînga trebuie să fie pozitiv, astfel încît cel puțin unul din numerele F_2, F_3, F_4, F_5 este pozitiv, ceea ce urmăream să arătăm.

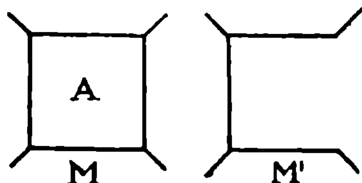


Fig. 146.

Acum să demonstrăm teorema celor cinci culori. Fie M o hartă regulată pe sferă, care conține n regiuni. Știm că cel puțin una dintre ele are mai puțin de șase laturi.

Cazul 1. M conține o regiune A cu două, trei sau patru laturi. În acest caz să îndepărtăm frontiera dintre A și una dintre regiunile adiacente. (Dacă A are patru laturi, una dintre regiuni poate atinge două laturi neadiacente ale lui A. În acest caz, din teorema lui Jordan, rezultă că regiunile care ating celelalte două laturi ale lui A vor fi distincte și atunci îndepărtăm frontiera dintre A și una dintre aceste regiuni.) Harta M′ care rezultă va fi o hartă regulată, formată din $n-1$ regiuni. *Dacă M′ poate fi colorată în mod potrivit cu cinci culori, atunci și* M *poate fi colorată în mod potrivit cu același număr de culori*, deoarece dacă cel mult patru regiuni ale lui M sînt adiacente cu A, putem găsi întotdeauna a cincea culoare pentru A.

Cazul 2. M conține o regiune A, cu cinci laturi. Să considerăm cele cinci regiuni adiacente lui A și să le numim B, C, D, E și F. Putem găsi întotdeauna o pereche dintre acestea care nu sînt adiacente pentru că, dacă de pildă, B și D s-ar atinge, atunci ele ar împiedica pe C să atingă pe E sau pe F, deoarece orice drum care duce de la C la E sau la F ar trebui să treacă cel puțin prin una dintre regiunile A, B și D (fig. 147). (Este limpede că și acest fapt depinde esențial de teorema lui Jordan, care este valabilă în plan sau pe

sferă. Ea nu este adevărată pe tor, de exemplu.) De aceea, putem presupune, de pildă, că C şi F nu se ating. Îndepărtăm laturile lui A, adiacente cu C şi F, formînd o nouă hartă M' cu $n - 2$ regiuni, care este şi ea regulată. *Dacă noua hartă poate fi colorată în mod potrivit cu cinci culori, atunci şi harta iniţială* M *poate fi colorată în mod potrivit cu acelaşi număr de culori.* Într-adevăr, dacă

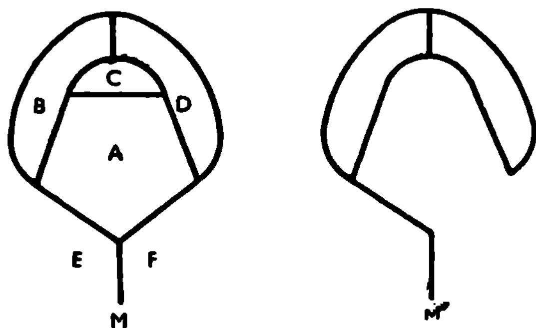


Fig. 147.

restabilim frontierele, A nu va atinge decît cel mult patru regiuni de culori diferite, deoarece C şi F avînd aceeaşi culoare, putem găsi o a cincea culoare pentru A.

Astfel, în orice caz, dacă M este o hartă regulată cu n regiuni, putem construi o nouă hartă regulată M', cu $n - 1$ sau $n - 2$ regiuni, astfel încît dacă M' poate fi colorată cu cinci culori, şi M poate fi colorată cu cinci culori. Acest procedeu poate fi aplicat din nou lui M' şi duce la un şir de hărţi deduse din M:

$$M, M', M'', \ldots.$$

Deoarece numărul regiunilor din hărţile acestui şir descreşte mereu, trebuie să ajungem în cele din urmă la o hartă cu cinci regiuni sau mai puţine. Dar o astfel de hartă poate fi colorată întotdeauna cu cel mult cinci culori, deci revenind la M pas cu pas, vedem că M poate fi colorată şi ea cu cinci culori. Aceasta încheie demonstraţia. Remarcaţi că această demonstraţie este constructivă pentru că ea dă o metodă practică, chiar dacă e plictisitoare, de colorare efectivă a oricărei hărţi, cu n regiuni, într-un număr finit de paşi.

2. Teorema lui Jordan pentru poligoane

Teorema lui Jordan afirmă că orice curbă simplă închisă C împarte punctele planului, care nu se află pe C, în două domenii distincte (fără puncte comune), pentru care C este frontiera comună. Vom da o demonstraţie pentru această teoremă, în cazul în care C este un *poligon* închis P.

Vom arăta că punctele din plan care nu se află pe *P* se împart în două clase *A* şi *B*, astfel încît două puncte ale aceleiaşi clase pot fi unite printr-un drum poligonal, care nu intersectează pe *P*, în timp ce orice drum care uneşte un punct al lui *A* cu un punct al lui *B* trebuie să intersecteze pe *P*. Clasa *A* va forma ,,exteriorul'' poligonului, în timp ce clasa *B* va forma ,,interiorul'' lui.

Începem demonstraţia alegînd o direcţie orientată fixă în plan, neparalelă cu nici una dintre laturile lui *P*. Deoarece *P* are un număr finit de laturi, acest lucru se poate face întotdeauna. Definim apoi clasele *A* şi *B* în modul următor:

Punctul p aparţine lui *A*, dacă semidreapta dusă prin p în direcţia orientată aleasă intersectează pe *P* într-un număr *par* 0, 2, 4, 6, ... de puncte. Punctul p aparţine clasei *B*, dacă semidreapta dusă prin p în direcţia orientată aleasă intersectează pe *P* într-un număr *impar* 1, 3, 5, ... de puncte.

Referitor la semidreptele care intersectează pe *P* în vîrfuri, vom conveni să nu numărăm o intersecţie într-un vîrf în care ambele laturi ale lui *P*, care pornesc din acel vîrf, sînt de aceeaşi parte a semidreptei, dar vom număra o intersecţie într-un vîrf în care cele două laturi sînt de o parte şi de alta a semidreptei. Vom spune că două puncte p şi q au aceeaşi ,,paritate'', dacă ele aparţin aceleiaşi clase *A* sau *B*.

Mai întîi observăm că toate punctele unui segment de dreaptă, care nu intersectează pe *P*, au aceeaşi paritate. Într-adevăr, paritatea unui punct p, care se deplasează în lungul unui astfel de segment, se poate schimba numai dacă semidreapta dusă prin p, în direcţia aleasă, trece printr-un vîrf al lui *P*, şi în nici unul dintre cele două cazuri posibile paritatea nu se va schimba,

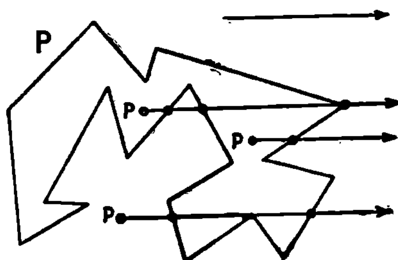


Fig. 148. Numărarea intersecţiilor

datorită convenţiei făcute mai sus. De aici rezultă că *dacă un punct* p_1 *al lui* A *se uneşte cu un punct* p_2 *al lui* B *printr-o linie poligonală, atunci acest drum trebuie să intersecteze pe* P, pentru că, în caz contrar, paritatea tuturor punctelor drumului, şi în particular a lui p_1 şi p_2, ar fi aceeaşi. Mai mult, putem arăta că *oricare două puncte din aceeaşi clasă* A *sau* B *pot fi unite printr-o linie poligonală care nu intersectează pe* P. Să notăm cele două puncte cu p şi q. Dacă segmentul rectiliniu pq, care uneşte pe p cu q, nu intersectează pe *P*, el este drumul dorit. În caz contrar, fie p' primul punct de intersecţie al acestui

segment cu P, şi fie q' ultimul punct de acest fel (fig. 149). Să construim drumul care porneşte din p în lungul segmentului pp', cotind apoi cu puţin înainte de p; mergînd paralel cu P, ajungem într-un punct de pe pq apropiat de q'. Dacă putem demonstra că acest drum va intersecta pe pq între q' şi q, şi nu între p' şi q', atunci drumul poate fi continuat pînă în q, în lungul lui $q'q$, fără a intersecta pe P. Este limpede că două puncte oarecare r şi s, destul de

Fig. 149.

apropiate, dar de o parte şi de alta a unui segment al lui P, trebuie să aibă paritate diferită, pentru că semidreapta dusă prin r va intersecta pe P într-un punct în plus faţă de semidreapta dusă prin s. Astfel vedem că paritatea se modifică atunci cînd trecem prin punctul q' în lungul segmentului pq. Rezultă că drumul punctat traversează segmentul pq între q' şi q, pentru că p şi q (şi deci orice punct al drumului punctat) au aceeaşi paritate.

Aceasta completează demonstraţia teoremei lui Jordan, pentru cazul unui poligon P. „Exteriorul" lui P poate fi identificat acum cu clasa A, deoarece dacă mergem destul de departe în lungul unei semidrepte duse în direcţia aleasă, vom ajunge într-un punct dincolo de care nu vom avea nici o intersecţie cu P, astfel încît toate aceste puncte au paritatea 0, deci aparţin lui A. Aceasta face ca „interiorul" lui P să fie identificat cu clasa B. Oricît de întortocheat ar fi poligonul simplu închis P, putem determina întotdeauna dacă un punct p dat din plan este interior sau exterior lui P, ducînd o semidreaptă şi numărînd intersecţiile semidreptei cu P. Dacă acest număr este impar, atunci punctul p este cuprins în interiorul lui P şi nu poate ieşi din el fără a traversa pe P printr-un punct. Dacă numărul este par, atunci punctul p este exterior lui P. (Verificaţi acest lucru pentru fig. 128.)

* Mai putem demonstra teorema lui Jordan pentru poligoane în modul următor: Să definim *ordinul* unui punct p_0 faţă de o curbă închisă C, care nu trece prin p_0, ca fiind numărul net de rotaţii complete, efectuate de o săgeată care uneşte pe p_0 cu un punct p mobil pe curbă, atunci cînd p parcurge curba o singură dată. Fie

A = toate punctele p_0 care nu se află pe P şi care au un ordin *par* în raport cu P,

B = toate punctele p_0 care nu se află pe P şi care au un ordin *impar* în raport cu P.

Atunci A şi B astfel definite formează exteriorul şi interiorul lui P. Efectuarea amănuntelor demonstraţiei este lăsată pe seama cititorului, ca exerciţiu.

*3. Teorema fundamentală a algebrei

„Teorema fundamentală a algebrei” afirmă că dacă

(1) $$f(z) = z^n + a_{n-1}z^{n-1} + a_{n-2}z^{n-2} + \ldots + a_1 z + a_0,$$

unde $n \geqslant 1$, și $a_{n-1}, a_{n-2}, \ldots, a_0$ sînt numere complexe, atunci există un număr complex α, astfel încît $f(\alpha) = 0$. Cu alte cuvinte, *în cîmpul numerelor complexe, orice ecuație polinomială are o rădăcină.* (La p. 118 am tras concluzia că $f(z)$ poate fi descompus sub forma unui produs de n factori liniari:

$$f(z) = (z - \alpha_1)(z - \alpha_2) \ldots (z - \alpha_n),$$

unde $\alpha_1, \alpha_2, \ldots, \alpha_n$ sînt zerourile lui $f(z)$.) Este remarcabil faptul că această teoremă poate fi demonstrată prin considerații cu caracter topologic, similare cu acelea folosite la demonstrarea teoremei de punct fix a lui Brouwer.

Cititorul își va reaminti că un număr complex este un simbol $x + yi$, unde x și y sînt numere reale, iar i are proprietatea că $i^2 = -1$. Numărul complex $x + yi$ poate fi reprezentat prin punctul din plan ale cărui coordonate, în raport cu o pereche de axe perpendiculare, sînt x, y. Dacă introducem în acest plan coordonatele polare, luînd originea și sensul pozitiv al axei Ox, respectiv ca pol și origine a unghiurilor, putem scrie

$$z = x + yi = r(\cos\theta + i\sin\theta),$$

unde $r = \sqrt{x^2 + y^2}$. Din formula lui De Moivre rezultă că

$$z^n = r^n(\cos n\theta + i \sin n\theta)$$

(cf.p. 113). Astfel, dacă facem ca numărul complex z să descrie un cerc de rază r cu centrul în origine, z^n va efectua n rotații complete în lungul cercului de rază r^n, atunci cînd z descrie cercul o singură dată. Reamintim, de asemenea, că r, modulul lui z, care se mai scrie și sub forma $|z|$, dă distanța dintre z și 0 și că dacă $z' = x' + iy'$, atunci $|z - z'|$ este distanța dintre z și z'. Cu aceste preliminarii putem trece la demonstrarea teoremei.

Să presupunem că polinomul (1) nu are nici o rădăcină, astfel încît pentru orice număr complex z avem

$$f(z) \neq 0.$$

Din această ipoteză rezultă că dacă facem pe z să descrie o curbă închisă în planul x, y, atunci $f(z)$ va descrie o curbă închisă Γ, care nu trece prin origine (fig. 150). De aceea, putem defini *ordinul* originii O față de funcția $f(z)$, pentru orice curbă închisă C, ca fiind *numărul net de rotații complete efectuate de o săgeată care unește pe* 0 *cu un punct de pe curba* Γ, *descrisă de punctul care reprezintă pe* $f(z)$, atunci cînd z descrie curba C. Drept curbă C vom lua un cerc cu centrul în origine și de rază t și vom defini funcția $\varphi(t)$, ca fiind

ordinul lui O față de funcția $f(z)$, pentru cercul cu centrul în O și cu raza t. Avem, evident $\varphi(0) = 0$, deoarece un cerc de rază O este un punct și curba Γ se reduce la punctul $f(0) \neq O$. Vom arăta mai jos că $\varphi(t) = n$ pentru valori mari ale lui t. Dar ordinul $\varphi(t)$ depinde continuu de t, deoarece $f(z)$ este o funcție continuă de z. Prin urmare vom avea o contradicție, pentru că funcția $\varphi(t)$ poate lua numai valori întregi și de aceea nu poate trece în mod continuu de la valoarea 0 la valoarea n.

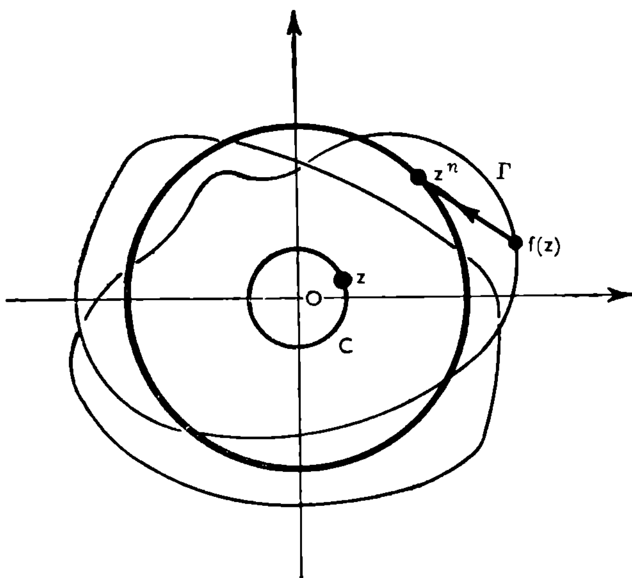


Fig. 150. Demonstrarea teoremei fundamentale a algebrei

Rămîne doar să arătăm că $\varphi(t) = n$ pentru vlori mari aleui t. Să observăm că pe un cerc de rază $z = t$, suficient de mare încît

$$t > 1 \text{ și } t > |a_0| + |a_1| + \ldots + |a_{n-1}|,$$

avem inegalitatea

$$|f(z) - z^n| = |a_{n-1}z^{n-1} + \ldots + a_0| \leqslant |a_{n-1}| \cdot |z|^{n-1} + $$
$$+ |a_{n-2}| \cdot |z|^{n-2} + \ldots + |a_0| =$$
$$= t^{n-1}\left[|a_{n-1}| + \ldots + \frac{|a_0|}{t^{n-1}}\right] \leqslant t^{n-1}[|a_{n-1}| + |a_{n-2}| + \ldots +$$
$$+ |a_0|] < t^n = |z^n|.$$

Deoarece expresia din stînga este distanţa dintre punctele z^n şi $f(z)$, în timp ce expresia din dreapta este distanţa de la punctul z^n la origine, vedem că segmentul de dreaptă care uneşte punctele $f(z)$ şi z^n nu poate trece prin origine, atîta timp cît z se află pe cercul de rază t, cu centrul în origine. Astfel stînd lucrurile, putem deforma în mod continuu curba descrisă de $f(z)$, transformînd-o în curba descrisă de z^n, fără a traversa originea, deplasînd pur şi simplu fiecare din punctele $f(z)$ în lungul segmentului care îl uneşte cu z^n. Deoarece ordinul originii va varia în mod continuu şi poate lua doar valori întregi în cursul acestei deformări, el trebuie să fie acelaşi pentru cele două curbe. Deoarece ordinul lui z^n este n, ordinul lui $f(z)$ trebuie să fie şi el egal cu n, ceea ce completează demonstraţia.

Capitolul VI

FUNCȚII ȘI LIMITE

INTRODUCERE

Cea mai importantă parte a matematicii moderne este centrată în jurul noțiunilor de funcție și limită. În acest capitol vom analiza sistematic aceste noțiuni.

O expresie, ca de pildă

$$x^2 + 2x - 3,$$

nu are o valoare numerică precisă, decît dacă este atribuită o valoare lui x. Spunem că valoarea acestei expresii este o *funcție* de valoarea lui x și scriem

$$x^2 + 2x - 3 = f(x).$$

De exemplu, dacă $x = 2$, atunci $2^2 + 2 \cdot 2 - 3 = 5$, astfel încît $f(2) = 5$. În același mod, prin substituția directă, putem găsi valoarea lui $f(x)$ pentru orice număr x întreg, rațional, irațional sau chiar complex.

Numărul numerelor prime mai mici decît n este o funcție $\pi(n)$ de întregul n. Dacă este dată o valoare a lui n, atunci valoarea lui $\pi(n)$ este determinată, chiar dacă nu este cunoscută nici o expresie algebrică pentru calcularea ei. Aria unui triunghi este funcție de lungimile celor trei laturi ale lui ; ea variază dacă lungimile laturilor se modifică și este determinată atunci cînd aceste lungimi au valori determinate. Dacă un plan este supus unei transformări proiective sau topologice, atunci coordonatele unui punct după transformare depind, adică sînt funcții, de coordonatele punctului inițial. Noțiunea de funcție intervine ori de cîte ori anumite cantități sînt legate printr-o anumită relație fizică. Volumul unui gaz cuprins într-un cilindru este funcție de temperatură și de presiunea pistonului. Presiunea atmosferică măsurată dintr-un balon este o funcție de altitudinea lui deasupra nivelului mării. Întregul domeniu al fenomenelor periodice — mișcarea mareelor, vibrațiile unei corzi ciupite, emisiunea undelor de lumină dintr-un filament incandescent — este guvernat de funcțiile trigonometrice simple sin x și cos x.

Pentru Leibniz (1646—1716), care a folosit pentru prima dată cuvîntul „funcţie", şi pentru matematicienii secolului al XVIII-lea, noţiunea de dependenţă funcţională era identificată mai mult sau mai puţin cu existenţa unei formule matematice simple, care să exprime natura exactă a acestei dependenţe. Această noţiune s-a dovedit a fi prea restrictivă pentru necesităţile fizicii matematice şi ideea de funcţie, împreună cu noţiunea înrudită de limită, a fost supusă unui lung proces de generalizare şi clarificare, despre care vom vorbi în acest capitol.

§ 1. VARIABILĂ ŞI FUNCŢIE

1. Definiţii şi exemple

Adesea avem de-a face cu obiecte matematice pe care le putem alege după voie dintr-o mulţime S de obiecte. Obiectul ales în acest caz îl numim *variabilă* în *domeniul* S. De obicei se folosesc litere de la sfîrşitul alfabetului pentru desemnarea variabilelor. Astfel, dacă S este mulţimea tuturor întregilor, variabila X din domeniul S reprezintă un înteg arbitrar. Spunem că „variabila X parcurge mulţimea S", înţelegînd prin aceasta că sîntem liberi să identificăm simbolul X cu orice element al mulţimii S. Folosirea noţiunii de variabilă este avantajoasă dacă dorim să facem afirmaţii referitoare la obiecte care pot fi arbitrar alese dintr-o mulţime. De exemplu, dacă S este din nou mulţimea întregilor, iar X şi Y sînt variabile din domeniul S, atunci propoziţia

$$X + Y = Y + X$$

este o exprimare simbolică convenabilă a faptului că suma a doi întregi arbitrari este independentă de ordinea în care ei sînt luaţi. Un caz particular este exprimat de egalitatea

$$2 + 3 = 3 + 2,$$

în care intervin constaute, dar pentru a exprima legea generală, valabilă pentru toate perechile de numere, sînt necesare simboluri care au înţelesul de variabile.

Nu este nicidecum necesar ca domeniul S al unei variabile X să fie o mulţime de numere. De exemplu, S poate fi mulţimea tuturor cercurilor din plan; atunci X va desemna un anumit cerc. Sau S ar putea fi mulţimea tuturor poligoanelor închise din plan, iar X un anumit poligon. De asemenea nu este necesar ca domeniul uuei variabile să conţină o infinitate de elemente. De exemplu, X ar putea desemua orice membru al populaţiei S a unui anumit oraş, la un anumit moment. Sau X ar putea desemna unul dintre resturile împărţirii unui întreg prin 5; în acest caz, domeniul S ar consta din cele cinci numere 0, 1, 2, 3, 4.

Cel mai important caz al unei variabile numerice — şi în acest caz folosim de obicei o literă mică x — este acela în care domeniul de variaţie S este un interval $a \leqslant x \leqslant b$ al axei numerice. Spunem atunci că x este o *variabilă continuă* în intervalul considerat. Domeniul de variaţie al unei variabile continue se poate întinde la infinit. Astfel, S poate fi mulţimea tuturor numerelor reale pozitive $x > 0$ sau chiar mulţimea tuturor numerelor reale. Într-un mod similar putem considera o variabilă X, ale cărei valori sînt punctele unui plan sau ale unui anumit domeniu din plan, ca de pildă interiorul unui dreptunghi sau al unui cerc. Deoarece orice punct al planului este definit prin cele două coordonate x, y în raport cu o pereche fixată de axe, spunem adesea că în acest caz avem de-a face cu o *pereche de variabile continue* x *şi* y.

Se poate întîmpla ca oricărei valori a unei variabile X să i se asocieze o anumită valoare a unei alte variabile U. Atunci U se numeşte *funcţie* de X. Modul în care U este legat de X se exprimă prin simboluri, ca de pildă

$$U = F(X) \qquad \text{(se citeşte „}F\text{ de }X\text{")}.$$

Dacă X parcurge mulţimea S, atunci variabila U va parcurge o altă mulţime T. De exemplu, dacă S este mulţimea tuturor triunghiurilor X din plan, se poate defini o funcţie $F(X)$ asociind fiecărui triunghi X lungimea $U = F(X)$ a perimetrului său; în acest caz, T va fi mulţimea tuturor numerelor pozitive. Observăm aici că două triunghiuri diferite X_1 şi X_2 pot avea acelaşi perimetru, astfel încît este posibilă egalitatea $F(X_1) = F(X_2)$, chiar dacă $X_1 \neq X_2$. O transformare proiectivă a unui plan S pe un altul, T, asociază fiecărui punct X al lui S un singur punct U al lui T, potrivit unei reguli bine determinate, pe care o putem exprima prin simbolul funcţional $U = F(X)$. În acest caz, $F(X_1) \neq F(X_2)$, ori de cîte ori $X_1 \neq X_2$ şi spunem că aplicaţia lui S pe T este *biunivocă* (cf. p. 95).

Funcţiile de variabilă continuă sînt definite adesea prin expresii algebrice, aşa cum sînt definite, de exemplu, următoarele funcţii

$$u = x^2, \qquad u = \frac{1}{x}, \qquad u = \frac{1}{1 + x^2}.$$

În prima şi ultima dintre aceste expresii, x poate parcurge întreaga mulţime a numerelor reale, în timp ce în a doua x poate parcurge mulţimea numerelor reale, cu excepţia lui 0 (valoarea 0 este exclusă, deoarece simbolul 1/0 nu este număr).

Numărul $B(n)$ al factorilor primi ai lui n este o funcţie de n, unde n parcurge domeniul tuturor numerelor naturale. Mai general, orice şir de numere $a_1, a_2, a_3, \ldots$ poate fi privit ca mulţime de valori ale unei funcţii $u = F(n)$, în care domeniul variabilei independente n este mulţimea numerelor naturale. Doar pentru a prescurta scrierea, vom nota al n-lea termen al şirului prin

a_n, în loc de a folosi notația funcțională mai explicită $F(n)$. Expresiile discutate în cap. I.

$$S_1(n) = 1 + 2 + \cdots + n = \frac{n(n+1)}{2},$$

$$S_2(n) = 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6},$$

$$S_3(n) = 1^3 + 2^3 + \cdots + n^3 = \frac{n^2(n+1)^2}{4},$$

sînt funcții de variabila întreagă n.

Dacă $U = F(X)$, folosim pentru X numele de *variabilă independentă*, în timp ce U este numită *variabila dependentă*, pentru că valoarea ei depinde de valoarea aleasă pentru X.

Se poate întîmpla ca aceeași valoare a lui U să fie atribuită tuturor valorilor lui X, astfel încît mulțimea T constă dintr-un singur element. Atunci avem cazul particular în care valoarea U a funcției nu variază, adică U este *constantă*. Vom include acest caz în noțiunea generală de funcție, chiar dacă aceasta ar putea să pară ciudat începătorului, pentru care accentul pare să cadă pe faptul că U variază atunci cînd variază X. Însă nu va fi nici o pagubă — și de fapt va fi chiar folositor — să privim o constantă ca fiind un caz particular de variabilă, al cărei „domeniu de variație" constă dintr-un singur element.

Noțiunea de funcție este de cea mai mare importanță nu numai în matematica pură, dar și în aplicațiile practice. Legile fizice nu sînt decît propoziții referitoare la modul în care anumite cantități depind de altele, atunci cînd acestea pot varia. Astfel, înălțimea notei emise de o coardă ciupită depinde de lungimea, greutatea și tensiunea corzii, presiunea atmosferică depinde de altitudine, iar energia unui glonț depinde de masa și de viteza lui. Misiunea fizicianului este să determine natura exactă sau aproximativă a acestei dependențe funcționale.

Noțiunea de funcție permite o caracterizare matematică exactă a mișcării. Dacă o particulă mobilă este concentrată într-un punct din spațiu, cu coordonatele rectangulare x, y, z și dacă t măsoară timpul, atunci mișcarea particulei este descrisă complet, prin indicarea coordonatelor x, y, z ca funcții de t:

$$x = f(t), \quad y = g(t), \quad z = h(t).$$

Astfel, dacă o particulă cade liber în lungul axei verticale Oz, aflată numai sub influența gravitației, atunci

$$x = 0, \quad y = 0, \quad z = -\frac{1}{2}gt^2,$$

unde g este accelerația gravitației. Dacă o particulă se rotește uniform pe un cerc de rază unitate din planul x, y, atunci mișcarea ei este caracterizată prin funcțiile

$$x = \cos \omega t, \quad y = \sin \omega t,$$

unde ω este o constantă, așa-numită viteză unghiulară a mișcării.

O funcție matematică este pur și simplu o lege care guvernează interdependența unor cantități variabile. Ea nu implică existența nici unei relații de tipul „cauză și efect" între ele. Cu toate că în limbajul obișnuit, cuvîntul „funcție" este folosit adesea cu acest înțeles, vom evita interpretările filozofice de acest fel. De exemplu, legea lui Boyle pentru un gaz conținut într-un recipient la temperatură constantă afirmă că produsul dintre presiunea p și volumul v este o constantă c (a cărei valoare depinde însă de temperatură):

$$pv = c.$$

Această ecuație poate fi rezolvată atît în raport cu p, cît și în raport cu v, ca funcție de cealaltă variabilă,

$$p = \frac{c}{v} \quad \text{sau} \quad v = \frac{c}{p},$$

fără ca aceasta să implice faptul că o schimbare a volumului este „cauza" unei schimbări a presiunii sau că modificarea presiunii este „cauza" modificării volumului. Pentru matematician prezintă interes numai forma legăturii dintre cele două variabile.

De remarcat că noțiunea de funcție este oarecum diferit abordată de matematicieni față de fizicieni. Primii subliniază de obicei *legea de corespondență*, operația matematică care este aplicată variabilei independente pentru a obține valoarea variabilei dependente u. În acest sens, $f(\ )$ este simbolul unei operații matematice; valoarea $u = f(x)$ este rezultatul aplicării operației $f(\ )$ numărului x. Pe de altă parte, fizicianul este adesea mai interesat în *cantitatea* u, ca atare, decît în procedeul matematic prin care valorile lui u pot fi calculate din acelea ale lui x. Astfel, rezistența u opusă de aer unui obiect în mișcare depinde de viteza v și poate fi obținută experimental, chiar dacă nu se cunoaște nici o formulă matematică explicită pentru calculul lui $u = f(v)$. Fizicianul este interesat în cunoașterea rezistenței reale și nu în cunoașterea formulei matematice particulare $f(v)$, cu excepția cazului în care studiul unei astfel de formule poate ajuta la analizarea comportării cantității u. Aceasta este atitudinea adoptată de obicei de cel care *aplică* matematica în fizică sau tehnică. În calculele mai avansate cu funcții, uneori se poate evita confuzia numai dacă știm exact dacă este vorba despre operația $f(\ )$, care asociază lui x o cantitate $u = f(x)$, sau despre cantitatea u însăși, care mai poate fi considerată că depinde într-un alt mod de o altă variabilă z. De exemplu, aria unui cerc este dată de funcția $u = f(x) = \pi x^2$, unde x este raza, dar mai este dată și de funcția $u = g(z) = z^2/4\pi$, unde z este circumferința.

Poate că cele mai simple tipuri de funcții matematice de o variabilă sînt *polinoamele*, de forma

$$u = f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

cu „coeficienți" constanți $a_0, a_1, \ldots, a_n$. După ele urmează *funcțiile raționale*, ca de pildă

$$u = \frac{1}{x}, \quad u = \frac{1}{1+x^2}, \quad u = \frac{2x+1}{x^4+3x^2+5},$$

care sînt cîturi de polinoame, și apoi *funcțiile trigonometrice* cos x, sin x și tg x = sin x/cos x, care pot fi definite cel mai bine cu ajutorul cercului unitate din planul ξ, η, de ecuație $\xi^2 + \eta^2 = 1$. Dacă punctul $P(\xi, \eta)$ descrie circumferința acestui cerc și dacă x este unghiul orientat, cu care trebuie să rotim semiaxa pozitivă $O\xi$, pentru a o face să coincidă cu OP, atunci cos x și sin x sînt coordonatele lui P: cos $x = \xi$, sin $x = \eta$.

2. Măsura unghiurilor în radiani

În toate aplicațiile practice, unghiurile se măsoară în unități obținute prin subdivizarea unui unghi drept într-un număr de părți egale. Dacă acest număr este 90, atunci unitatea este cunoscutul „grad". O subdivizare în 100 de părți egale ar fi mai potrivită pentru sistemul nostru zecimal, însă ar reprezenta același principiu de măsurare. Pentru scopuri teoretice însă, este avantajos să folosim o metodă esențial diferită de caracterizare a mărimii unui unghi, așa-numita măsurare prin radiani. Multe formule importante, în care intervin funcțiile trigonometrice ale unghiurilor, au o formă mai simplă în acest sistem, în comparație cu acela în care unghiurile se măsoară prin grade.

Pentru a găsi măsura în radiani a unui unghi, descriem un cerc de rază 1 cu centrul în vîrful unghiului. Unghiul va intercepta un arc s pe circumferința cercului și luăm prin definiție lungimea acestui arc ca *măsură în radiani* a unghiului. Deoarece întreaga circumferință a cercului de rază 1 are lungimea 2π, întregul unghi de 360° are măsura de 2π radiani. Rezultă că dacă x este măsura în radiani a unui unghi, iar y este măsura în grade a aceluiași unghi, atunci x și y sînt legați prin relația $y/360 = x/2\pi$ sau

$$\pi y = 180\, x.$$

Astfel, un unghi de 90°($y = 90$) are măsura $x = 90\,\pi/180 = \pi/2$ radiani etc. Pe de altă parte, un unghi de 1 radian (unghiul a cărui măsură în radian este $x = 1$) este unghiul care interceptează un arc egal cu raza cercului; în grade, acesta va fi un unghi de $y = 180/\pi = 57{,}2957\ldots$ grade. Trebuie să înmulțim întotdeauna măsura în radiani x a unui unghi cu factorul $180/\pi$ pentru a obține măsura lui în grade y.

Măsura în radiani x a unui unghi oarecare este egală cu dublul ariei A a sectorului cercului unitate determinat de unghiul dat, pentru că această arie se află față de aria întregului cerc în același raport în care arcul interceptat pe circumferință se află față de întreaga circumferință: $x/2\pi = A/\pi$, $x = 2A$.

De acum înainte, unghiul x va fi unghiul a cărui măsură în radiani este egală cu x. Un unghi de x grade va fi scris $x°$, pentru a evita ambiguitatea.

Va fi evident că măsura în radiani este foarte convenabilă pentru operațiile analitice, însă pentru scopuri practice va fi mai degrabă neconvenabilă. Deoarece π este irațional, nu vom reveni niciodată în același punct al cercului, dacă purtăm succesiv unghiul unitate, adică unghiul a cărui măsură în radiani este egală cu 1. Măsura obișnuită este astfel concepută, încît după purtarea unui grad de 360 de ori sau a cîte 90 de grade de patru ori, să revenim la poziția inițială.

3. Graficul unei funcții. Funcții inverse

Natura unei funcții devine evidentă de cele mai multe ori cu ajutorul unui grafic geometric simplu. Dacă x, u sînt coordonatele în plan în raport cu o pereche de axe perpendiculare, atunci funcțiile liniare de forma

$$u = ax + b$$

sînt reprezentate prin linii drepte; funcțiile pătratice, ca de pildă

$$u = ax^2 + bx + c,$$

prin parabole; funcția

$$u = \frac{1}{x}$$

printr-o hiperbolă etc. Prin definiție, *graficul* oricărei funcții $u = f(x)$ constă din toate punctele din plan, ale căror coordonate x, u sînt legate prin relația $u = f(x)$. Funcțiile $\sin x$, $\cos x$, $\operatorname{tg} x$ sînt reprezentate prin curbele din fig. 151 și 152. Aceste grafice arată limpede cum crește sau descrește valoarea funcției atunci cînd x variază.

O metodă importantă de introducere a unor noi funcții este următoarea. Începînd cu o funcție cunoscută $F(X)$, putem încerca să rezolvăm ecuația $U = F(X)$ în raport cu X, astfel încît X să apară ca funcție de U:

$$X = G(U).$$

Funcția $G(U)$ se numește atunci *funcția inversă* a lui $F(X)$. Acest procedeu duce la un rezultat unic, numai dacă funcția $U = F(X)$ definește o aplicație biunivocă a domeniului lui X pe acela al lui U, adică dacă inegalitatea $X_1 \neq X_2$

implică întotdeauna inegalitatea $F(X_1) \neq F(X_2)$, deoarece numai atunci vom avea un X unic determinat, asociat fiecărui U. Exemplul precedent, în care X desemna un triunghi din plan, iar $U = F(X)$ era perimetrul său, se pretează la o discuție. Evident, această aplicație a mulțimii S a triunghiurilor pe mulțimea T a numerelor reale pozitive nu este biunivocă, deoarece există o infinitate de triunghiuri diferite care au acelaşi perimetru. Deci, în acest

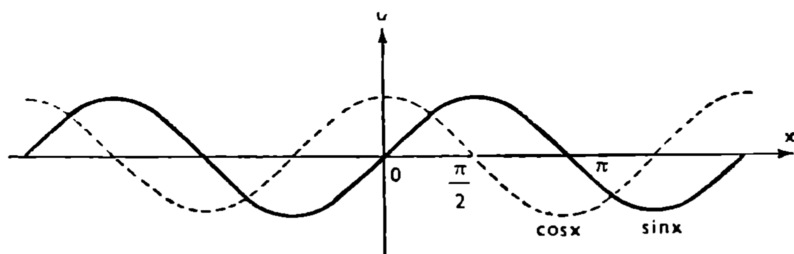


Fig. 151. Graficul lui sin x şi cos x

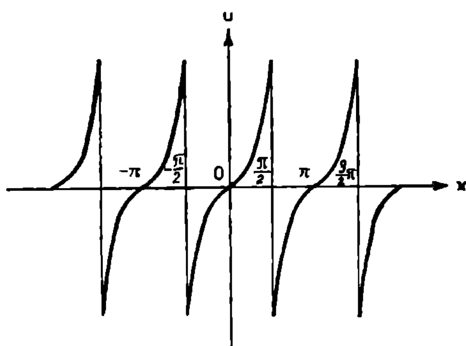


Fig. 152. u = *tg* x

caz, relația $U = F(X)$ nu poate defini o funcție inversă uniformă. Pe de altă parte, funcția $m = 2n$, în care n parcurge mulțimea S a întregilor iar m mulțimea T a întregilor pari, dă o corespondență biunivocă între cele două mulțimi şi funcția inversă $n = m/2$ este definită în mod unic. Un alt exemplu de aplicație biunivocă este dat de funcția

$$u = x^3.$$

Cînd x parcurge mulțimea tuturor numerelor reale, u va parcurge , de asemenea, mulțimea tuturor numerelor reale, luînd fiecare valoare o singură dată. Funcția inversă, definită în mod unic, este

$$x = \sqrt[3]{u}.$$

În cazul funcției

$$u = x^2,$$

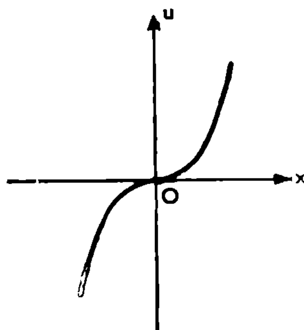


Fig. 153. $u = x^3$

funcția inversă nu este determinată în mod unic. Într-adevăr, din $u = x^2 = (-x)^2$, rezultă că orice valoare pozitivă a lui u are *două* preimagini. Dar dacă, așa cum se face de obicei, înțelegem prin simbolul $\sqrt{u}$, numărul *pozitiv* al cărui pătrat este egal cu u, atunci funcția inversă

$$x = \sqrt{u}$$

există, atîta timp, cît x și u reprezintă valori pozitive.

Existența unei inverse unice a unei funcții de o variabilă $u = f(x)$ poate fi stabilită din examinarea graficului funcției. Funcția inversă va fi definită în mod unic numai dacă fiecărei valori a lui u îi corespunde o singură valoare a lui x. Utilizînd graficul, aceasta înseamnă că nici o paralelă la axa Ox nu intersectează graficul în mai mult de un punct. Aceasta se va întîmpla cu siguranță, dacă funcția $u = f(x)$ este *strict monotonă*, adică dacă funcția este strict crescătoare sau strict descrescătoare, atunci cînd x crește. De exemplu, dacă $u = f(x)$ este strict crescătoare, atunci pentru $x_1 < x_2$, avem întotdeauna $u_1 = f(x_1) < u_2 = f(x_2)$. Deci pentru o valoare dată a lui u, există cel mult un x, astfel încît $u = f(x)$, și funcția inversă va fi determinată în mod unic. Graficul funcției inverse $x = g(u)$ se obține prin simpla rotire

a graficului inițial cu un unghi de 180°, în jurul dreptei trasate punctat (fig. 154), astfel încît pozițiile axei Ox și axei Ou sînt schimbate între ele. Noua poziție a graficului va da pe x ca funcție de u. În poziția inițială, graficul dă pe u ca înălțime deasupra axei orizontale Ox, în timp ce după rotație, același grafic dă pe x ca înălțime deasupra axei orizontale Ou.

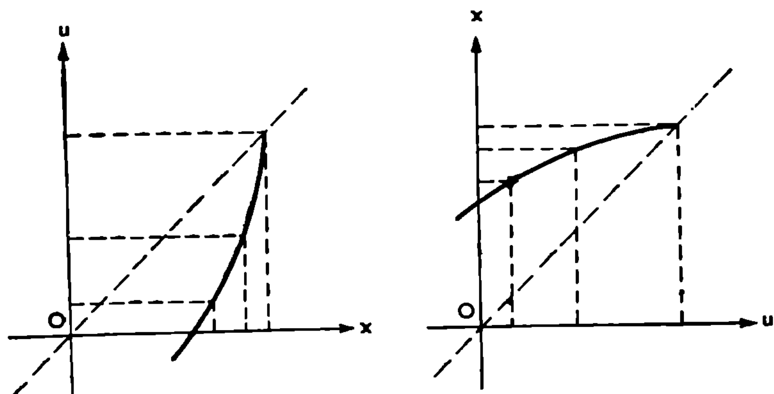


Fig. 154. Funcții inverse

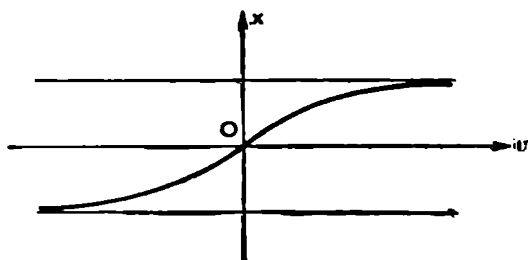


Fig. 155. $x = \text{arc tg } u$

Considerațiile precedente pot fi ilustrate pentru cazul funcției

$$u = \operatorname{tg} x.$$

Această funcție este strict monotonă pentru $-\pi/2 < x < \pi/2$ (fig. 152). Valorile lui u, care cresc strict o dată cu x, parcurg intervalul de la $-\infty$ la $+\infty$; deci funcția inversă

$$x = g(u)$$

este definită pentru toate valorile lui u. Această funcție se notează prin $\text{arc tg } u$. Astfel, $\text{arc tg}(1) = \pi/4$, deoarece $\operatorname{tg} \pi/4 = 1$. Graficul ei este reprezentat în fig. 155.

4. Funcţii compuse

O altă metodă importantă de formare a unor noi funcţii din două sau mai multe funcţii date este *compunerea* funcţiilor. De exemplu, funcţia

$$u = f(x) = \sqrt{1 + x^2}$$

este „compusă" din două funcţii mai simple

$$z = g(x) = 1 + x^2, \qquad u = h(z) = \sqrt{z},$$

şi poate fi scrisă ca

$$u = f(x) = h(g(x)) \quad \text{(se citeşte „}h \text{ de } g \text{ de } x\text{")}.$$

De asemenea

$$u = f(x) = \frac{1}{\sqrt{1 - x^2}}$$

este compusă din următoarele trei funcţii

$$z = g(x) = 1 - x^2, \quad w = h(z) = \sqrt{z}, \quad u = k(w) = \frac{1}{w},$$

astfel încît

$$u = f(x) = k(h(g(x))).$$

Funcţia

$$u = f(x) = \sin \frac{1}{x}$$

este compusă din următoarele două funcţii

$$z = g(x) = \frac{1}{x}, \; u = h(z) = \sin z.$$

Funcţia $f(x)$ nu este definită pentru $x = 0$, deoarece pentru $x = 0$ expresia $1/x$ nu are sens. Graficul acestei funcţii remarcabile se obţine din acela al funcţiei sinus. Ştim că $\sin z = 0$ pentru $z = k\pi$, unde k este orice întreg pozitiv sau negativ. Mai mult,

$$\sin z = \begin{cases} 1 \text{ pentru } z = (4k + 1)\dfrac{\pi}{2} \\ -1 \text{ pentru } z = (4k - 1)\dfrac{\pi}{2} \end{cases},$$

dacă k este un întreg oarecare. Deci

$$\sin\frac{1}{x}=\begin{cases}0 \text{ pentru } x=\dfrac{1}{k\pi},\\ 1 \text{ pentru } x=\dfrac{2}{(4k+1)\pi},\\ -1 \text{ pentru } x=\dfrac{2}{(4k-1)\pi}.\end{cases}$$

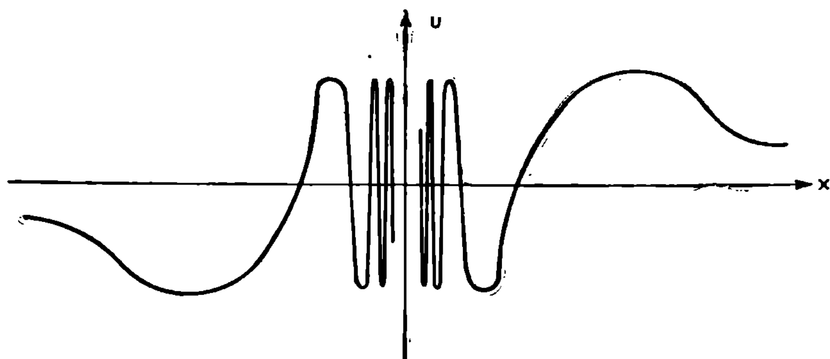


Fig. 156. u = *sin* 1/x

Dacă punem succesiv

$$k = 1, 2, 3, 4, \ldots,$$

atunci, deoarece numitorii acestor fracții cresc nelimitat, valorile lui x pentru care funcția $\sin 1/x$ are valorile 1, -1, 0, se vor acumula în punctul $x = 0$. Între orice punct de acest fel și origine va exista o infinitate de oscilații ale funcției. Graficul funcției este arătat în fig. 156.

5. Continuitatea

Graficele funcțiilor considerate pînă acum dau o imagine intuitivă despre continuitate. Vom face o analiză precisă a acestei noțiuni în § 4, după ce noțiunea de limită va fi pusă pe o bază riguroasă. Pe scurt, spunem că o funcție este continuă, dacă graficul ei este o curbă neîntreruptă (cf. p. 328). O funcție dată $u = f(x)$ poate fi verificată în privința continuității, făcînd ca variabila independentă x să varieze continuu dinspre dreapta și dinspre stînga, spre orice valoare aleasă x_1. Dacă funcția $u = f(x)$ nu este constantă în vecinătatea lui x_1, valoarea ei se va schimba. Dacă valoarea lui $f(x)$ are ca limită valoarea $f(x_1)$ a funcției în punctul ales $x = x_1$, *oricum ne-am apropia de* x_1, *dintr-o parte sau din alta*, atunci se spune că funcția este *continuă* în x_1. Dacă acest lucru

se întîmplă pentru orice punct x_1 al unui anumit interval, atunci se spune că funcția este ***continuă pe acest interval.***

Cu toate că orice funcție reprezentată printr-un grafic nerupt este continuă, este ușor să definim funcții care nu sînt continue peste tot. De exemplu, funcția din fig. 157, definită pentru toate valorile lui x prin

$$f(x) = \quad 1 + x \text{ pentru } x > 0,$$
$$f(x) = -1 + x \text{ pentru } x \leqslant 0,$$

este discontinuă în punctul $x_1 = 0$, unde are valoarea -1. Dacă încercăm să trasăm graficul acestei funcții, va trebui să ridicăm creionul de pe hîrtie în acest punct. Dacă ne apropiem de valoarea $x_1 = 0$ dinspre dreapta, atunci $f(x)$ tinde spre $+1$. Dar această valoare diferă de valoarea -1 a funcției din acest punct. Faptul că -1 este valoarea către care tinde $f(x)$, atunci cînd x tinde către 0 dinspre *stînga*, nu este suficient pentru stabilirea continuității.

Funcția $f(x)$, definită pentru orice x prin

$$f(x) = 0 \quad \text{pentru } x \neq 0, \quad f(0) = 1,$$

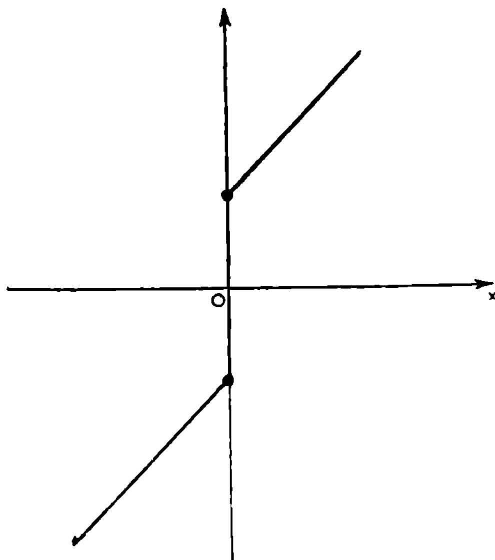


Fig. 157. Discontinuitate prin salt

are o discontinuitate de natură diferită în punctul $x_1 = 0$. Aici, atît limita la dreapta, cît și cea de la stînga există și sînt egale cu 0, atunci cînd x tinde spre 0, dar această valoare comună a limitelor este diferită de $f(0)$.

Un alt tip de discontinuitate este arătat de funcția din fig. 158,

$$u = f(x) = \frac{1}{x^2},$$

în punctul $x = 0$. Dacă facem pe x să tindă către zero din ambele părți, atunci u tinde spre infinit; graficul funcției este întrerupt în acest punct, și variații mici ale lui x în vecinătatea lui $x = 0$ pot produce variații foarte mari ale

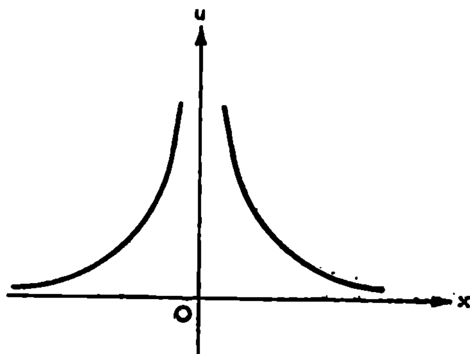


Fig. 158. Discontinuitate prin infinit

lui u. Riguros vorbind, valoarea funcției nu este definită pentru $x = 0$, pentru că nu admitem infinitul ca număr și de aceea nu putem spune că $f(x)$ *este egală* cu infinit atunci cînd $x = 0$. Deci spunem doar că $f(x)$ „tinde către infinit", atunci cînd x tinde spre zero.

Un alt tip de discontinuitate apare în cazul funcției $u = \sin(1/x)$ în punctul $x = 0$, așa cum se vede din graficul acestei funcții (fig. 156).

Exemplele precedente indică mai multe moduri în care o funcție poate să nu fie continuă într-un punct $x = x_1$:

1) Este posibil să facem funcția continuă în $x = x_1$, definind sau redefinind în mod potrivit valoarea ei pentru $x = x_1$. De exemplu, funcția $u = x/x$ este mereu egală cu 1, cînd $x \neq 0$; ea nu este definită pentru $x = 0$, deoarece $0/0$ este un simbol fără sens. Dacă însă în acest caz convenim ca valoarea $u = 1$ să corespundă și valorii $x = 0$, atunci funcția astfel prelungită devine continuă pentru orice valoare a lui x, fără excepție. Obținem același rezultat dacă redefinim funcția definită la pagina precedentă punînd $f(0) = 0$. O discontinuitate de acest fel poartă numele de discontinuitate *aparentă*.

2) Cînd x tinde spre x_1 dinspre dreapta și dinspre stînga, funcția poate avea limite diferite, ca în fig. 157.

3) Este posibil ca limitele laterale să nu existe, ca în fig. 156.

4) Funcția poate tinde spre infinit, atunci cînd x tinde spre x_1, ca în fig. 158.

Discontinuitățile de aceste trei tipuri se numesc *esențiale*; ele nu pot fi eliminate, definind sau redefinind în mod potrivit funcția doar în punctul $x = x_1$.

Exerciții: 1) Reprezentați funcțiile $\frac{x-1}{x^2}$, $\frac{x^2-1}{x^2+1}$, $\frac{x}{(x^2-1)(x^2+1)}$ și găsiți discontinuitățile lor.

2) Reprezentați funcțiile $x \sin(1/x)$ și $x^2 \sin(1/x)$ și verificați faptul că ele sînt continue în $x = 0$, dacă punem în ambele cazuri $u = 0$ pentru $x = 0$.

*3) Arătați că funcția arc tg $(1/x)$ are o discontinuitate de al doilea tip (prin salt) pentru $x = 0$.

*6. Funcții de mai multe variabile

Să revenim la discutarea sistematică a noțiunii de funcție. Dacă variabila independentă P este un punct din plan, cu coordonatele x, y, și dacă fiecărui punct P îi corespunde un singur număr u (de exemplu, u ar putea fi distanța de la punctul P la origine), atunci de obicei scriem

$$u = f(x,y).$$

Această notație se folosește și dacă, așa cum se întîmplă adesea, cele două cantități x și y apar de la început ca variabile independente. De exemplu, presiunea u a unui gaz este o funcție de volumul x și de temperatura y, iar aria u a unui triunghi este o funcție $u = f(x, y, z)$ de lungimile x, y și z ale celor trei laturi.

În același mod în care graficul dă o reprezentare geometrică a unei funcții de o variabilă, reprezentarea geometrică a unei funcții $u = f(x, y)$ de două variabile este dată de o suprafață din spațiul tridimensional, cu coordonatele x, y, u. Fiecărui punct x, y din planul x, y îi asociem un punct în spațiu, ale cărui coordonate sînt x, y și $u = f(x,y)$. Astfel, funcția $u = \sqrt{1 - x^2 - y^2}$ este reprezentată printr-o suprafață sferică de ecuație $u^2 + x^2 + y^2 = 1$, funcția liniară $u = ax + by + c$ printr-un plan, funcția $u = xy$ prin paraboloidul hiperbolic etc.

Putem da o altă reprezentare a funcției $u = f(x,y)$, în planul x, y, cu ajutorul *curbelor de nivel*. În loc de a considera „peisajul" tridimensional al funcției $u = f(x,y)$, trasăm ca pe o hartă curbele de nivel ale funcției, indicînd proiecțiile pe planul x, y ale tuturor punctelor care au aceeași cotă u. Aceste curbe de nivel sînt pur și simplu curbele $f(x,y) = c$, unde c rămîne constantă pentru fiecare curbă. Astfel, funcția $u = x + y$ este caracterizată de fig. 163. Curbele de nivel ale unei suprafețe sferice sînt cercuri concentrice.

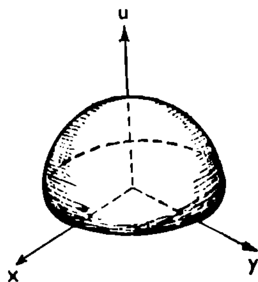


Fig. 159. Emisfera

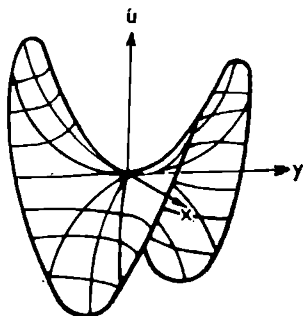


Fig. 160. Paraboloidul hiperbolic

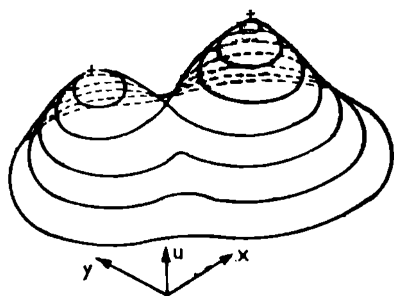


Fig. 161. Suprafață de forma u = f(x,y)

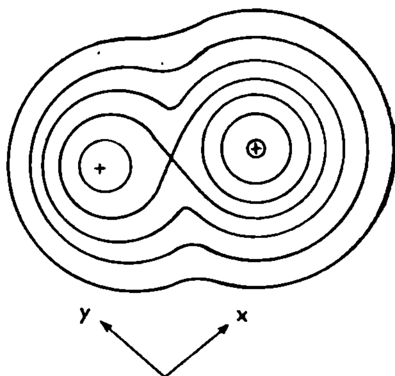


Fig. 162. Curbele de nivel ale suprafeței reprezentate în fig. 161

Funcția $u = x^2 + y^2$, care reprezintă un paraboloid de rotație, este caracterizată, de asemenea, prin cercuri (fig. 165). Ca număr atașat oricărei curbe putem indica cota corespunzătoare $u = c$.

Funcțiile de mai multe variabile apar în fizică atunci cînd trebuie descrisă mișcarea unui mediu continuu. De exemplu, să presupunem că o coardă este întinsă între două puncte de pe axa Ox și apoi este deformată, astfel încît particula cu poziția x este deplasată perpendicular pe axă. Dacă dăm apoi drumul corzii, ea va vibra, astfel încît particula care inițial avea coordonata x se va afla în momentul t la o distanță $u = f(x,t)$ de axa Ox. Mișcarea este descrisă complet prin cunoașterea funcției $u = f(x,t)$.

Definiția continuității dată pentru funcții de o singură variabilă se extinde imediat la funcțiile de mai multe variabile. O funcție $u = f(x,y)$ este conti-

nuă în punctul $x = x_1$, $y = y_1$, dacă $f(x,y)$ tinde spre $f(x_1, y_1)$ atunci cînd punctul x, y tinde spre punctul x_1, y_1, din orice direcţie şi în orice mod.

Există totuşi o deosebire importantă între funcţiile de o variabilă şi cele de mai multe variabile. În ultimul caz, noţiunea de funcţie inversă devine fără sens, deoarece nu putem rezolva o ecuaţie $u = f(x, y)$, ca de pildă

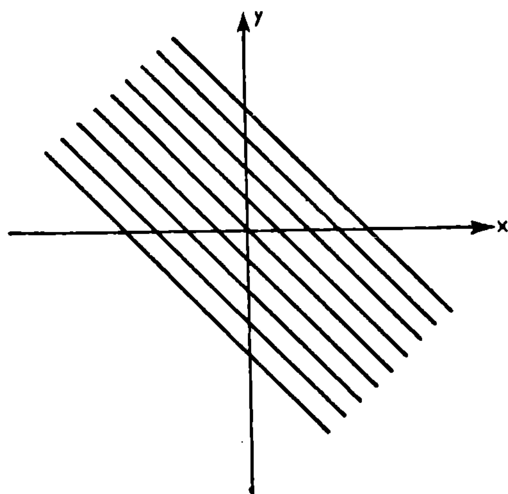


Fig. 163. Curbele de nivel ale lui u = x + y

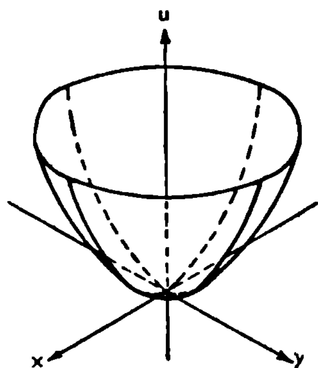


Fig. 164. Paraboloidul de revoluţie

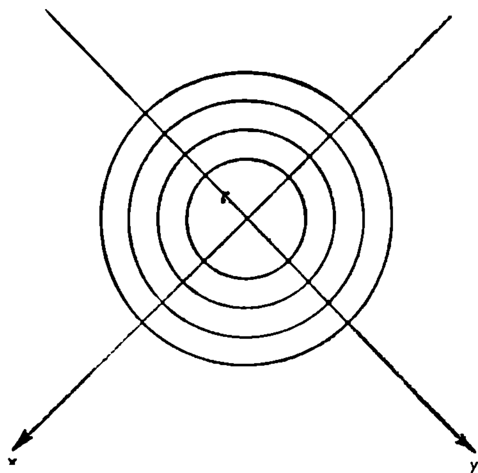


Fig. 165. Curbele de nivel ale suprafeţei din fig. 164

$u = x + y$, astfel încît fiecare din cantitățile independente x, y să poată fi exprimată numai în funcție de *singura* cantitate u. Dar această deosebire dintre funcțiile de o variabilă și cele de mai multe variabile dispare, dacă subliniem faptul că o funcție definește o aplicație sau o transformare.

*7. Funcții și transformări

O corespondență dintre punctele unei drepte l, caracterizată printr-o coordonată x în lungul dreptei, și punctele unei alte drepte l', caracterizată printr-o coordonată x', este pur și simplu o funcție $x' = f(x)$. Dacă corespondența este biunivocă, avem și o funcție inversă $x = g(x')$. Cel mai simplu exemplu este transformarea prin proiecție, care — și afirmă acest lucru fără demonstrație — este caracterizată în general printr-o funcție de forma $x' = f(x) = (ax + b)/(cx + d)$, unde a, b, c, d sînt constante. În acest caz, funcția inversă este $x = g(x') = (-dx' + b)/(cx' - a)$.

Aplicațiile în două dimensiuni, de la un plan π, cu coordonatele x, y, pe un plan π', cu coordonatele x', y', nu pot fi reprezentate printr-o singură funcție $x' = f(x)$, ci necesită două funcții de două variabile:

$$x' = f(x, y),$$

$$y' = g(x, y).$$

De exemplu, o transformare proiectivă este dată de un sistem de funcții,

$$x' = \frac{ax + by + c}{gx + hy + k},$$

$$y' = \frac{dx + ey + f}{gx + hy + k},$$

unde a, b, ..., k sînt constante și unde x, y și x', y' sînt coordonatele din cele două plane. Din acest punct de vedere, ideea de transformare inversă are sens. Trebuie să *rezolvăm* doar *acest sistem de ecuații* în raport cu x și y, în funcție de x' și y'. Din punct de vedere geometric, aceasta se reduce la găsirea aplicației inverse de la planul π' pe planul π. Aceasta va fi definită în mod unic, dacă corespondența dintre punctele celor două plane este biunivocă.

Transformările planului studiate în topologie nu sînt date prin ecuații algebrice simple, ci prin orice sistem de funcții,

$$x' = f(x,y),$$

$$y' = g(x,y),$$

care definesc o transformare biunivocă și bicontinuă.

Exerciții: *1) Arătați că inversiunea (cap. III, p. 159) în raport cu cercul unitate este dată analitic prin ecuațiile $x' = x/(x^2 + y^2)$, $y' = y/(x^2 + y^2)$. Găsiți transformarea inversă. Demonstrați analitic că inversiunea transformă dreptele și cercurile în drepte sau cercuri.

2) Demonstrați că prin transformarea $x' = (ax + b)/(cx + d)$ patru puncte de pe axa Ox sînt transformate în patru puncte de pe axa Ox', care au același biraport (cf. p. 194).

§ 2. LIMITE

1. Limita unui șir a_n

Așa cum am văzut în § 1, descrierea continuității unei funcții este bazată pe noțiunea de limită. Pînă acum am folosit această noțiune sub o formă mai mult sau mai puțin intuitivă. În cele ce urmează, o vom studia într-un mod mai sistematic. Deoarece șirurile sînt mai simple decît funcțiile de o variabilă continuă, vom începe cu studiul șirurilor.

În cap. II am întîlnit șiruri a_n de numere și am studiat limitele lor, atunci cînd n „tinde spre infinit". De exemplu, șirul al cărui termen de indice n este $a_n = 1/n$,

$$1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots, \tag{1}$$

are limita 0, cînd n tinde spre infinit:

$$\frac{1}{n} \to 0, \text{ cînd } n \to \infty. \tag{2}$$

Să încercăm să enunțăm precis ce înțelegem prin aceasta. Pe măsură ce mergem din ce în ce mai departe în șir, termenii devin din ce în ce mai mici. După cel de-al 100-lea termen, toți termenii sînt mai mici decît 1/100, după cel de-al 1 000-lea termen, toți termenii sînt mai mici decît 1/1 000 și așa mai departe. Nici unul dintre termeni nu este chiar egal cu zero, însă dacă mergem *destul de departe* în șirul (1), putem fi siguri că fiecare dintre termenii lui va diferi de 0 cu *oricît de puțin vrem.*

Singura supărare pricinuită de această explicație este faptul că înțelesul frazelor subliniate nu este suficient de clar. Cît de departe este „destul de departe", și cît de mic este „oricît de mic vrem"? Dacă putem asocia un înțeles precis acestor fraze, atunci putem da un înțeles precis relației (2).

O interpretare geometrică ne va ajuta să facem situația mai clară. Dacă reprezentăm termenii șirului (1) prin punctele corespunzătoare de pe axa numerică, atunci observăm că termenii șirului par să se acumuleze spre punctul 0. Să alegem un interval I oarecare de pe axa numerică, cu centrul în punctul 0 și de lungime totală 2ε, astfel încît intervalul se întinde pe o distanță egală

cu ε de fiecare parte a punctului 0. Dacă alegem $\varepsilon = 10$, atunci, desigur, *toţi* termenii $a_n = 1/n$ ai şirului se vor afla în interiorul intervalului I. Dacă alegem $\varepsilon = 1/10$, atunci primii cîţiva termeni ai şirului se vor afla în exteriorul lui I, însă toţi termenii, începînd cu a_{11},

$$\frac{1}{11}, \frac{1}{12}, \frac{1}{13}, \frac{1}{14}, \ldots,$$

se vor afla în interiorul lui I. Chiar dacă alegem $\varepsilon = 1/1\,000$, doar primii 1 000 de termeni ai şirului nu se vor afla în interiorul lui I, în timp ce toţi termenii, începînd cu a_{1001},

$$a_{1001}, a_{1002}, a_{1003}, \ldots$$

se vor afla în interiorul lui I. Desigur, acest raţionament se poate face pentru orice număr ε pozitiv: îndată ce am ales un ε pozitiv oricît de mic ar fi, putem găsi un întreg N, astfel încît

$$\frac{1}{N} < \varepsilon.$$

De aici rezultă că toţi termenii a_n ai şirului, pentru care $n \geqslant N$, se vor afla în interiorul lui I, şi numai termenii $a_1, a_2, \ldots, a_{N-1}$ se pot afla în exterior. Faptul important este următorul: *în primul rînd*, lungimea intervalului I este aleasă după voie, prin alegerea lui ε. *Apoi*, poate fi găsit un întreg N potrivit. Acest proces, care constă în alegerea unui număr ε şi apoi în găsirea unui întreg N potrivit, poate fi efectuat pentru orice număr pozitiv ε, oricît de mic, şi dă un înţeles precis propoziţiei că toţi termenii şirului (1) vor diferi de 0, cu oricît de puţin dorim, cu condiţia să mergem destul de departe în şir.

Să rezumăm: fie ε un număr pozitiv arbitrar. Atunci putem găsi un întreg N, astfel încît toţi termenii a_n din şirul (1), pentru care $n \geqslant N$, se vor afla în interiorul intervalului I, de lungime totală 2ε, cu centrul în punctul O. Acesta este înţelesul precis al relaţiei (2).

Pe baza acestui exemplu, putem da acum o definiţie precisă a propoziţiei generale: „şirul de numere reale $a_1, a_2, a_3, \ldots$ are limita a". Includem pe a în interiorul unui interval I de pe axa numerică: dacă intervalul este mic, unii termeni a_n se pot afla în exteriorul intervalului, însă îndată ce n devine suficient de mare, de pildă mai mare sau egal cu un întreg N, atunci toate numerele a_n pentru care $n \geqslant N$ trebuie să se afle în interiorul intervalului I. Desigur, s-ar putea să fie necesar ca întregul N să fie ales foarte mare, dacă a fost ales un interval I foarte mic; însă oricît de mic ar fi intervalul I, trebuie să existe un astfel de întreg N pentru ca şirul să aibă limita a.

Faptul că un şir a_n are limita a se exprimă simbolic prin notaţia

$$\lim a_n = a \quad \text{cînd } n \to \infty,$$

sau mai simplu,

$$a_n \to a \quad \text{cînd } n \to \infty$$

(se citeşte: a_n *tinde spre* a, *sau converge către* a). Definiţia convergenţei unui şir a_n către a poate fi formulată mai concis în modul următor: *şirul* a_1, a_2, a_3, ... *are limita* a, *cînd* n *tinde spre infinit, dacă oricărui număr pozitiv* ε, *oricît de mic, i se poate asocia un întreg* N (*care depinde de* ε), *astfel încît*:

(3) $$|a - a_n| < \varepsilon$$

pentru orice

$$n \geqslant N.$$

Aceasta este formularea abstractă a noţiunii de limită a unui şir. Nu este de mirare faptul că la prima confruntare cu această noţiune, ea nu poate fi înţeleasă în cîteva minute. Există o atitudine nefericită, aproape snoabă, din partea unor autori de manuale, care prezintă cititorului această definiţie fără o pregătire prealabilă, ca şi cum explicarea nu ar fi de demnitatea unui matematician.

Definiţia sugerează o întrecere între două persoane, A şi B. Persoana A pretinde ca o cantitate fixată a să fie aproximată prin a_n cu o precizie mai mare decît o margine aleasă $\varepsilon = \varepsilon_1$; B satisface pretenţia, demonstrînd că există un întreg $N = N_1$, astfel încît toţi termenii a_n, care urmează după elementul a_{N_1}, satisfac această condiţie. Atunci A poate deveni mai exigent şi va fixa o nouă margine mai mică, $\varepsilon = \varepsilon_2 \cdot B$ satisface din nou pretenţia lui A, găsind un întreg $N = N_2$ (eventual mult mai mare). *Dacă* B *poate mulţumi pe* A, *oricît de mică ar fi marginea aleasă de* A, *atunci avem situaţia exprimată de* $a_n \to a$.

Există o anumită dificultate psihologică în înţelegerea acestei definiţii precise a limitei. Intuiţia noastră ne sugerează o idee „dinamică" despre noţiunea de limită, ca rezultat al unui proces de „mişcare": ne deplasăm în lungul şirului de întregi 1, 2, 3, ..., n, ... şi atunci observăm comportarea şirului a_n. Ni se pare că am putea observa faptul că $a_n \to a$. Însă această atitudine „naturală" nu este capabilă de o formulare matematică, riguroasă. Pentru a ajunge la o definiţie precisă, trebuie să *inversăm* ordinea paşilor; în loc de a privi mai întîi la variabila independentă n şi apoi la variabila dependentă a_n, trebuie să ne bazăm definiţia noastră pe ceea ce trebuie să facem dacă vrem să verificăm afirmaţia că $a_n \to a$. Printr-un astfel de procedeu, trebuie să alegem mai întîi o mică margine arbitrară în jurul lui a şi apoi trebuie să stabilim dacă putem satisface această condiţie, luînd variabila independentă n suficient de mare. Apoi, dînd denumiri simbolice, ε şi N frazelor „margine arbitrar de mică" şi „n suficient de mare", sîntem conduşi la definiţia precisă a limitei.

Ca un alt exemplu, să considerăm şirul

$$\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots, \frac{n}{n+1}, \ldots,$$

unde $a_n = \dfrac{n}{n+1}$. Afirm că $\lim a_n = 1$. Dacă alegeţi un interval cu centrul în punctul 1 şi pentru care $\varepsilon = 1/10$, atunci pot satisface pretenţia dv. (3), alegînd $N = 10$, pentru că

$$0 < 1 - \frac{n}{n+1} = \frac{n+1-n}{n+1} = \frac{1}{n+1} < \frac{1}{10}$$

îndată ce $n \geqslant 10$. Dacă vă măriţi pretenţia, alegînd $\varepsilon = 1/1000$, atunci o pot satisface din nou, alegînd $N = 1000$; şi în mod similar, pentru orice număr pozitiv ε, oricît de mic, pe care l-aţi putea alege; de fapt, trebuie să aleg doar un întreg N mai mare decît $1/\varepsilon$. Acest proces de fixare a unei margini ε arbitrar de mică în jurul numărului a şi de demonstrare a faptului că termenii şirului a_n se află toţi la o distanţă mai mică decît ε faţă de a, dacă mergem destul de departe în şir, este descrierea detaliată a faptului că $\lim a_n = a$.

Dacă termenii şirului $a_1, a_2, a_3, \ldots$ sînt exprimaţi ca fracţii zecimale infinite, atunci afirmaţia $\lim a_n = a$ înseamnă pur şi simplu că pentru orice întreg pozitiv m, primele m zecimale ale lui a_n coincid cu primele m zecimale ale dezvoltării în fracţie zecimală infinită a numărului fixat a, cu condiţia ca n să fie ales suficient de mare, de pildă mai mare sau egal cu o valoare N (care depinde de m). Aceasta corespunde alegerii unui ε de forma 10^{-m}.

Mai există o altă cale, foarte sugestivă, de exprimare a noţiunii de limită. Dacă $\lim a_n = a$ şi dacă includem pe a în interiorul unui interval I, atunci oricît de mic ar fi I, toate numerele a_n pentru care n este mai mare sau egal cu un întreg N se vor afla în interiorul lui I, astfel încît cel mult *un număr finit*, N — 1, *de termeni* de la începutul şirului,

$$a_1, a_2, \ldots, a_{N-1},$$

se pot afla în exteriorul lui I. Dacă I este foarte mic, N poate fi foarte mare, de pildă 100 sau chiar 1 000 de miliarde; totuşi, doar un număr finit de termeni ai şirului se vor afla în exteriorul lui I, în timp ce o infinitate de termeni se vor afla în interiorul lui I.

Despre termenii unui şir infinit putem spune că ,,aproape toţi'' au o anumită proprietate, dacă numai un număr finit de termeni, care poate fi oricît de mare, nu au acea proprietate. De pildă, ,,aproape toţi'' întregii pozitivi sînt mai mari decît 1 000 000 000 000. Folosind această terminologie, afirmaţia $\lim a_n = a$ este echivalentă cu următoarea afirmaţie: *dacă*

I *este un interval cu centrul în* a, *atunci aproape toate numerele* a_n *se află în interiorul lui* I.

Ar trebui să remarcăm în treacăt că nu se presupune că toți termenii a_n ai șirului au valori diferite. Este permis ca o infinitate dintre ei, sau chiar *toți* termenii a_n să fie egali cu valoarea limită a. De exemplu, șirul pentru care $a_1 = 0,\ a_2 = 0,\ \ldots,\ a_n = 0,\ \ldots$ este un șir legitim și desigur limita lui este 0.

Un șir a_n cu o limită a se numește *convergent*. Un șir a_n fără limită se numește *divergent*.

Exerciții : Demonstrați că :

1) Șirul pentru care $a_n = \dfrac{n}{n^2+1}$ are limita 0. (Indicație : $a_n = \dfrac{1}{n + \dfrac{1}{n}}$ este mai mic decît $1/n$ și mai mare decît 0.)

2) Șirul $a_n = \dfrac{n^2+1}{n^3+1}$ are limita 0. $\left(\text{Indicație : } a_n = \dfrac{1+\dfrac{1}{n^2}}{n+\dfrac{1}{n^2}} \text{ se află între 0 și } 2/n.\right)$

3) Șirul 1, 2, 3, 4, ... și șirurile oscilante

1, 2, 1, 2, 1, 2, ...

$-1, 1, -1, 1, -1, 1, \ldots$ (adică $a_n = (-1)^n$) și $1, \dfrac{1}{2}, 1, \dfrac{1}{3}, 1, \dfrac{1}{4}, \ldots$ *nu* au limită.

Dacă într-un șir a_n termenii devin atît de mari încît a_n este mai mare decît orice număr K, ales de la început, atunci spunem că a_n *tinde spre infinit* și scriem $\lim a_n = \infty$ sau $a_n \to \infty$. De exemplu, $n^2 \to \infty$ și $2^n \to \infty$. Această terminologie este folositoare, deși nu este chiar corectă, pentru că ∞ nu este un număr. *Un șir care tinde spre infinit este un șir divergent.*

Exercițiu : Demonstrați că șirul $a_n = \dfrac{n^2+1}{n}$ tinde spre infinit ; în mod asemănător pentru $a_n = \dfrac{n^2+1}{n+1}$, $a_n = \dfrac{n^3-1}{n+1}$ și $a_n = \dfrac{n^n}{n^2+1}$.

Uneori începătorii greșesc, gîndind că o trecere la limită cînd $n \to \infty$ poate fi efectuată, pur și simplu, făcînd substituția $n = \infty$ în expresia lui a_n. De exemplu, $1/n \to 0$, pentru că „$1/\infty = 0$”. Însă simbolul ∞ nu este un număr și folosirea lui în expresia $1/\infty$ nu este legitimă. Încercînd să ne imaginăm limita unui șir ca fiind „ultimul” termen a_n cînd $n = \infty$ pierdem esențialul și întunecăm situația.

2. Şiruri monotone

În definiția generală de la p.308 nu se cere un anumit tip de apropiere a unui şir convergent $a_1, a_2, a_3, \ldots$ spre limita sa a. Cel mai simplu tip este pus în evidență de aşa-numitul şir monoton, ca de pildă şirul

$$\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots, \frac{n}{n+1}, \ldots,$$

Fiecare termen al acestui şir este mai mare decît termenul precedent, deoarece $a_{n+1} = \frac{n+1}{n+2} = 1 - \frac{1}{n+2} > 1 - \frac{1}{n+1} = \frac{n}{n+1} = a_n$. Un şir de acest fel, în care $a_{n+1} > a_n$, se numeşte *strict crescător*. În mod asemănător, un şir pentru care $a_n > a_{n+1}$, ca de pildă şirul 1, 1/2, 1/3, ... se numeşte *strict descrescător*[1]. Aceste şiruri se pot apropia de limita lor numai dintr-o parte. Spre deosebire de acestea, există şiruri care oscilează, ca de pildă şirul $-1, +1/2, -1/3, +1/4, \ldots$ Acest şir se apropie de limita lui care este 0, din ambele părți (fig. 11).

Comportarea unui şir monoton este uşor de determinat. Un astfel de şir poate să nu aibă limită, şi atunci termenii lui se îndepărtează mereu, ca de pildă şirul

$$1, 2, 3, 4, \ldots,$$

în care $a_n = n$ sau şirul

$$2, 3, 5, 7, 11, 13, \ldots,$$

în care a_n este cel de-al n-lea număr prim p_n. În acest caz, şirul tinde spre infinit. Dar dacă termenii unui şir strict crescător rămîn mărginiți — adică dacă fiecare termen este mai mic decît o anumită margine B, cunoscută de la început —, atunci este clar, din punct de vedere intuitiv, că şirul trebuie să tindă spre o anumită limită a, care va fi mai mică sau cel mult egală cu B. Formulăm acest fapt sub forma *principiului şirurilor monotone*: *orice şir monoton crescător, care are o margine superioară, converge spre o limită.*

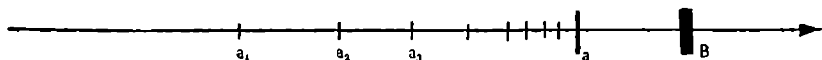


Fig. 166. Şir monoton şi mărginit

(O propoziție asemănătoare este valabilă pentru orice şir monoton *descrescător*, cu margine *inferioară*.) Este remarcabil faptul că valoarea limitei a nu trebuie să fie dată sau cunoscută de la început; teorema afirmă că, în con-

[1] Un şir a_n se numeşte monoton crescător, dacă $a_n \leqslant a_{n+1}$, pentru orice n; monoton descrescător, dacă $a_n \geqslant a_{n+1}$ pentru orice n. — *N.T.*

dițiile enunțate, limita *există*. Desigur, această teoremă depinde de introducerea numerelor iraționale și fără ele nu ar fi întotdeauna adevărată. Într-adevăr, așa cum am văzut în cap. II, orice număr irațional, (ca de pildă $\sqrt{2}$) este limita șirului mărginit și monoton crescător de fracții zecimale raționale, obținute prin secționarea unei anumite fracții zecimale infinite după cea de-a n-a zecimală.

*Cu toate că principiul șirurilor monotone apelează la intuiție, ca adevăr evident, va fi instructiv să dăm o demonstrație riguroasă sub o formă modernă. Pentru a face acest lucru, trebuie să arătăm că principiul este o consecință logică a definițiilor numărului real și a limitei.

Să presupunem că numerele $a_1, a_2, a_3, \ldots$ formează un șir monoton crescător și mărginit. Putem exprima termenii acestui șir sub forma unor fracții zecimale infinite,

$$a_1 = A_1, p_1p_2p_3 \ldots,$$

$$a_2 = A_2, q_1q_2q_3 \ldots,$$

$$a_3 = A_3, r_1r_2r_3 \ldots,$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

unde A_i sînt întregi, iar p_i, q_i, r_i etc. sînt zecimale cuprinse între 0 și 9. Acum, să coborîm în lungul coloanei întregilor $A_1, A_2, A_3, \ldots$ Deoarece șirul $a_1, a_2, a_3, \ldots$ este *mărginit*, acești întregi nu pot crește nemărginit, și deoarece șirul este *monoton crescător*, șirul de întregi A_1, $A_2, A_3, \ldots$ *va rămîne constant după atingerea valorii maxime.* Să notăm cu A această valoare maximă și să presupunem că ea este atinsă la linia de indice N_0. Acum, să coborîm în lungul celei de-a doua coloane $p_1, q_1, r_1, \ldots$ concentrîndu-ne atenția asupra termenilor cuprinși în linia de indice N_0 și asupra celor următori. Dacă x_1 este cea mai mare zecimală care apare în această coloană după linia de indice N_0, atunci x_1 va rămîne *constant* după prima sa apariție, pe care o putem presupune ca avînd loc în linia de indice N_1, unde $N_1 \geqslant N_0$. Într-adevăr, dacă zecimala din această coloană ar descrește, la un anumit moment șirul $a_1, a_2, a_3, \ldots$, nu ar fi monoton crescător. Apoi considerăm zecimalele $p_2, q_2, r_2, \ldots$ din cea de-a treia coloană. Un raționament asemănător arată că după un anumit întreg $N_2 \geqslant N_1$, zecimalele celei de-a treia coloane sînt mereu egale cu o anumită cifră x_2. Dacă repetăm acest raționament pentru cea de-a patra, a cincea ... coloană, obținem cifrele $x_3, x_4, x_5, \ldots$ și întregii corespunzători N_3, N_4, $N_5, \ldots$. Este ușor de văzut că numărul

$$a = A, x_1\, x_2\, x_3\, x_4 \ldots$$

este limita șirului $a_1, a_2, a_3, \ldots$, pentru că dacă ε este ales $\geqslant 10^{-m}$, atunci pentru orice $n \geqslant N_m$, partea întreagă și primele m cifre care urmează după virgula din a_n vor coincide cu acelea ale lui a, astfel încît diferența $|a - a_n|$ nu poate depăși pe 10^{-m}. Deoarece acest lucru poate fi făcut pentru orice ε pozitiv, oricît de mic, alegînd pe m suficient de mare, teorema este demonstrată.

Această teoremă mai poate fi demonstrată și pe baza oricăreia din celelalte definiții ale numărului real, date în cap. II; de exemplu, folosind definiția cu șiruri descrescătoare de intervale sau definiția cu tăieturi Dedekind. Astfel de demonstrații pot fi găsite în cele mai multe manuale de analiză.

Principiul şirurilor monotone ar fi putut fi folosit în cap. II pentru a defini suma şi produsul a două fracţii zecimale infinite pozitive,

$$a = A,\ a_1a_2a_3\ \ldots,$$

$$b = B,\ b_1b_2b_3\ \ldots.$$

Două astfel de expresii nu pot fi adunate sau înmulţite în modul obişnuit, începînd de la dreapta spre stînga, deoarece nu există capătul din dreapta (ca exemplu, cititorul ar putea încerca să adune fracţiile zecimale infinite 0,333333... şi 0,989898...). Dar dacă x_n este fracţia zecimală *finită*, obţinută prin secţionarea expresiilor lui a şi b după cea de-a n-a zecimală, şi adunate apoi ca de obicei, atunci şirul $x_1, x_2, x_3, \ldots$ va fi monoton crescător şi mărginit (de către întregul $A + B + 2$, de pildă). Deci acest şir are o limită şi putem defini $a + b = \lim x_n$. Un proces similar se foloseşte la definirea produsului ab. Aceste definiţii pot fi extinse apoi cu ajutorul regulilor obişnuite ale aritmeticii, pentru a fi aplicate tuturor cazurilor, în care a şi b sînt numere pozitive sau negative.

Exerciţiu: Arătaţi pe această cale că suma celor două fracţii zecimale infinite, considerate mai sus, este numărul real 1,323232... = 131/99.]

Importanţa noţiunii de limită în matematică constă în faptul că *multe numere sînt definite numai ca limite*, adesea ca limite de şiruri monotone şi mărginite. Acesta este motivul pentru care cîmpul numerelor raţionale, în care aceste limite pot să nu existe, este prea restrîns pentru nevoile matematicii.

3. Numărul *e* al lui Euler

Numărul e şi-a dobîndit un loc stabil în matematică, împreună cu numărul π al lui Arhimede, după publicarea cărţii lui Euler *Introductio in Analysin Infinitorum*, în 1748. Aici se ilustrează minunat modul în care se poate utiliza principiul şirurilor monotone pentru a defini un nou număr real. Folosind prescurtarea

$$n! = 1 \cdot 2 \cdot 3 \cdot 4 \ldots n$$

pentru produsul primilor n întregi, să considerăm şirul $a_1, a_2, a_3, \ldots$, unde

$$a_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}. \tag{4}$$

Termenii a_n formează un şir monoton şi strict crescător, deoarece a_{n+1} provine din a_n, prin adunarea unei valori pozitive $\frac{1}{(n+1)!}$. Mai mult, valorile lui a_n sînt mărginite superior:

$$a_n < B = 3. \tag{5}$$

Într-adevăr, avem $\frac{1}{s!} = \frac{1}{2} \cdot \frac{1}{3} \cdots \frac{1}{s} < \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} = \frac{1}{2^{s-1}}$, şi deci

$$a_n < 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots + \frac{1}{2^{n-1}} = 1 + \frac{1 - \left(\frac{1}{2}\right)^n}{1 - \frac{1}{2}} =$$

$$= 1 + 2\left(1 - \left(\frac{1}{2}\right)^n\right) < 3,$$

unde am folosit formula dată la p. 29 pentru suma primilor n termeni ai unei progresii geometrice. Deci, în virtutea principiului şirurilor monotone, a_n trebuie să se apropie de o limită cînd n tinde spre infinit; *să notăm cu* e *această limită*. Pentru a exprima faptul că $e = \lim a_n$, putem scrie pe e sub forma „seriei infinite”

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!} + \ldots \tag{6}$$

Această „egalitate”, cu un şir de puncte la sfîrşit, este doar un alt mod de exprimare a conţinutului afirmaţiilor

$$a_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}$$

şi

$$a_n \to e, \text{ cînd } n \to \infty.$$

Seria (6) permite calculul lui e, cu orice grad de precizie dorit. De exemplu, suma (cu 9 cifre) a termenilor din (6), pînă la $1/12!$ inclusiv, este $\Sigma = 2{,}71828182\ldots$ (Cititorul ar trebui să verifice acest rezultat.) „Eroarea”, adică diferenţa dintre această valoare şi adevărata valoare a lui e, poate fi apreciată cu uşurinţă. Pentru diferenţa $(e - \Sigma)$, avem expresia

$$\frac{1}{13!} + \frac{1}{14!} + \ldots < \frac{1}{13!}\left(1 + \frac{1}{13} + \frac{1}{13^2} + \ldots\right) = \frac{1}{13!} \cdot \frac{1}{1 - \frac{1}{13}} = \frac{1}{12 \cdot 12!}.$$

Aceasta este atît de mică, încît ea nu poate influenţa cea de-a noua cifră a lui Σ. Deci, ţinînd seama de eventuala eroare, posibilă în ultima cifră a valorii date mai sus, avem $e = 2{,}7182818\ldots$, cu opt cifre.

**Numărul* e *este iraţional.* Pentru a demonstra aceasta, vom proceda indirect, presupunînd că $e = p/q$, unde p şi q sînt întregi, deducînd apoi o absurditate din această ipoteză. Deoarece

ştim că $2 < e < 3$, e nu poate fi număr întreg şi de aceea q trebuie să fie cel puţin egal cu 2. Acum, înmulţim ambii membri ai lui (6) cu $q! = 1 \cdot 2 \cdot 3 \cdot \ldots q$, obţinînd

$$(7) \qquad e \cdot q! = p \cdot 2 \cdot 3 \ldots (q-1) = [q! + q! + 3 \cdot 4 \ldots q + 4 \cdot 5 \ldots q + \ldots + (q-1)q + q + 1] + \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \ldots .$$

În membrul stîng avem evident un întreg. În membrul drept, termenul din paranteză este şi el un întreg. Restul din membrul drept însă este un număr pozitiv mai mic decît $1/2$, şi deci nu este întreg. Deoarece $q \geqslant 2$, termenii seriei $1/(q+1) + \ldots$ sînt respectiv cel mult egali cu termenii corespunzători ai seriei geometrice $1/3 + 1/3^3 + 1/3^3 + \ldots$ a cărei sumă este $1/3[1/(1 - 1/3)] = 1/2$. Deci egalitatea (7) este contradictorie: întregul din stînga nu poate fi egal cu numărul din dreapta, deoarece acesta din urmă, fiind suma unui întreg cu un număr pozitiv mai mic decît $1/2$, nu poate fi întreg.

4. Numărul π

După cum se ştie din matematica de liceu, lungimea circumferinţei unui cerc de rază unitate poate fi definită ca limită a unui şir de perimetre ale unor poligoane regulate, cu un număr crescător de laturi. Lungimea circumferinţei astfel definită, se notează cu 2π. Mai precis, dacă p_n este lungimea poligonului regulat înscris cu n laturi, iar q_n este lungimea poligonului regulat circumscris cu n laturi, atunci $p_n < 2\pi < q_n$. Mai mult, atunci cînd n creşte, fiecare din şirurile p_n, q_n se apropie de 2π monoton, şi cu fiecare pas obţinem o margine mai mică a erorii efectuate prin aproximarea lui 2π prin numerele p_n sau q_n. La p. 141 am găsit expresia

$$p_{2^m} = 2^m \sqrt{2 - \sqrt{2 + \sqrt{2 + \ldots}}},$$

care conţine $m - 1$ radicali suprapuşi. Această formulă poate fi folosită pentru a calcula valoarea aproximativă a lui 2π.

Exerciţii: 1) Găsiţi valoarea aproximativă a lui π, dată de p_4, p_8 şi p_{16}.

*2) Găsiţi o formulă pentru q_{2^m}.

*3) Folosiţi această formulă pentru a găsi pe q_4, q_8 şi q_{14}. Cunoscînd pe p_{14} şi q_{16}, găsiţi marginile între care trebuie să se afle π.

Ce *este* numărul π? Inegalitatea $p_n < 2\pi < q_n$ dă răspunsul complet, indicînd un şir descrescător de intervale, care se contractă spre punctul 2π. Totuşi, acest răspuns lasă de dorit, pentru că nu ne dă nici o informaţie referitoare la natura lui π, ca număr real: este el raţional sau iraţional, algebric sau transcendent? După cum am menţionat la p. 157, π este de fapt un număr transcendent şi deci iraţional. Spre deosebire de demonstraţia pentru e, demonstraţia iraţionalităţii lui π, dată pentru prima dată de J. H. Lambert (1728—1777), este mai dificilă şi nu va fi dată aici. Însă o altă informaţie referitoare la π ne este accesibilă. Reamintindu-ne afirmaţia,

că întregii constituie materialul de bază al matematicii, ne putem întreba dacă numărul π se află într-o legătură simplă cu întregii. Dezvoltarea lui π în fracție zecimală, cu toate că a fost calculată cu cîteva sute de zecimale, nu dezvăluie nici o urmă de regularitate. Acest lucru nu este cîtuși de puțin surprinzător, deoarece π și 10 nu au nimic în comun. Totuși, în secolul al

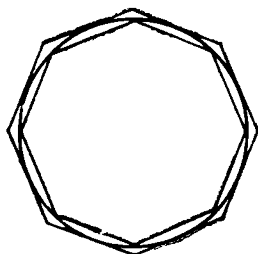


Fig. 167. Cercul aproximat prin poligoane

XVIII-lea, Euler și alții au găsit expresii frumoase, care leagă pe π de întregi, cu ajutorul seriilor și produselor infinite. Poate că cea mai simplă formulă de acest fel este următoarea:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots,$$

care exprimă pe $\pi/4$ ca limită a sumelor parțiale

$$s_n = 1 - \frac{1}{3} + \frac{1}{5} - \cdots + (-1)^n \frac{1}{2n+1},$$

cînd n crește. Vom obține această formulă în cap. VIII. O altă serie infinită pentru π este

$$\frac{\pi^2}{6} = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \dots.$$

O altă expresie surprinzătoare pentru π a fost descoperită de matematicianul englez John Wallis (1616—1703). Formula lui afirmă că

$$\left\{\frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdots \frac{2n}{2n-1} \cdot \frac{2n}{2n+1}\right\} \to \frac{\pi}{2} \text{ cînd } n \to \infty.$$

Uneori, aceasta se scrie sub forma prescurtată

$$\frac{\pi}{2} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdot \frac{8}{7} \cdot \frac{8}{9} \cdots$$

expresia din dreapta fiind numită un *produs infinit*.

O demonstrație a ultimelor două formule poate fi găsită în orice manual cuprinzător de analiză (cf. p. 502 și p. 530).

* 5. Fracții continue

Procese interesante de trecere la limită apar în legătură cu fracțiile continue. O fracție continuă finită, ca de pildă

$$\frac{57}{17} = 3 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{5}}},$$

reprezintă un număr rațional. La p. 65 am arătat că orice număr rațional poate fi scris sub această formă, cu ajutorul algoritmului lui Euclid. Pentru numerele iraționale însă, algoritmul nu se oprește după un număr finit de pași. În schimb, el duce la un șir de fracții de lungime din ce în ce mai mare, fiecare reprezentînd un număr rațional. În particular, toate numerele algebrice reale (cf. p. 120) de gradul doi pot fi exprimate în acest mod. Să considerăm, de exemplu, numărul $x = \sqrt{2} - 1$, care este o rădăcină a ecuației pătratice

$$x^2 + 2x = 1 \quad \text{sau} \quad x = \frac{1}{2 + x}.$$

Dacă înlocuim în membrul drept pe x prin $1/(2 + x)$, obținem expresia

$$x = \cfrac{1}{2 + \cfrac{1}{2 + x}},$$

și apoi

$$x = \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + x}}},$$

și așa mai departe, astfel încît după n pași, obținem egalitatea

$$\left. x = \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cdots + \cfrac{1}{2 + x}}}} \right\} n \text{ pași}.$$

Atunci cînd n tinde spre infinit, obținem „fracția continuă infinită"

$$\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cdots}}}.$$

Această formulă remarcabilă leagă pe $\sqrt{2}$ de întregi, într-un mod mai surprinzător decît dezvoltarea lui $\sqrt{2}$ în fracție zecimală, care nu manifestă nici o regularitate în succesiunea zecimalelor ei.

Pentru rădăcina pozitivă a oricărei ecuații pătratice de forma

$$x^2 = ax + 1 \quad \text{sau} \quad x = a + \frac{1}{x},$$

obținem dezvoltarea

$$x = a + \cfrac{1}{a + \cfrac{1}{a + \cfrac{1}{a + \cdots}}}.$$

De exemplu, punînd $a = 1$, găsim

$$x = \frac{1}{2}(1 + \sqrt{5}) = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}}$$

(cf. p. 139). Aceste exemple sînt cazuri particulare ale unei teoreme generale, care afirmă că *rădăcinile reale ale ecuațiilor pătratice cu coeficienți întregi*

admit dezvoltări în fracție continuă periodică, tot așa cum numerele raționale admit dezvoltări în fracții zecimale periodice.

Euler a reușit să găsească fracții continue infinite aproape tot atît de simple pentru e și pentru π. Următoarele egalități sînt date fără demonstrație:

$$e = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{6 + \cdots}}}}}}}};$$

$$e = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{2}{3 + \cfrac{3}{4 + \cfrac{4}{5 + \cdots}}}}};$$

$$\frac{\pi}{4} = \cfrac{1}{1 + \cfrac{1^2}{2 + \cfrac{3^2}{2 + \cfrac{5^2}{2 + \cfrac{7^2}{2 + \cfrac{9^2}{2 + \cdots}}}}}}.$$

§ 3. LIMITE PRIN APROPIERE CONTINUĂ

1. Introducere. Definiţie generală

În § 2, secţiunea 1, am reuşit să dăm o formulare precisă propoziţiei „şirul a_n (adică funcţia $a_n = F(n)$, de variabilă întreagă n) are limita a, cînd n tinde spre infinit". Vom da acum o definiţie corespunzătoare propoziţiei „funcţia $u = f(x)$, de variabilă continuă x, are limita a, cînd x tinde spre valoarea x_1". Sub o formă intuitivă, această noţiune de limită prin apropiere continuă a variabilei independente x a fost folosită în § 1, secţiunea 5, pentru verificarea continuităţii funcţiei $f(x)$.

Să începem din nou cu un exemplu particular. Funcţia $f(x) = \frac{(x + x^3)}{x}$ este definită pentru toate valorile lui x, diferite de $x = 0$, în care numitorul se anulează. Dacă trasăm un grafic al funcţiei $u = f(x)$ pentru valori ale lui x din vecinătatea lui 0, este evident că atunci cînd x „se apropie" de 0, din ambele părţi, valoarea corespunzătoare a lui $u = f(x)$ „se apropie" de limita 1. Pentru a da o descriere precisă a acestui fapt, să găsim o formulă explicită pentru diferenţa dintre valoarea $f(x)$ şi numărul fixat 1:

$$f(x) - 1 = \frac{x + x^3}{x} - 1 = \frac{x + x^3 - x}{x} = \frac{x^3}{x}.$$

Dacă convenim să considerăm doar valori ale lui x apropiate de 0, dar nu şi valoarea $x = 0$ (pentru care $f(x)$ nu este definită), putem împărţi atît numitorul, cît şi numărătorul expresiei din membrul drept al acestei egalităţi prin x, obţinînd formula mai simplă

$$f(x) - 1 = x^2.$$

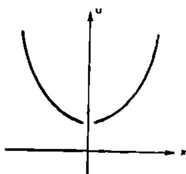


Fig. 168. u = (x + x³)/x

Desigur, putem face această diferenţă *oricît de mică vrem*, constrîngînd pe x să rămînă într-o vecinătate *suficient de mică* a valorii 0. Astfel, pentru $x = \pm 1/10$, $f(x) - 1 = 1/100$; pentru $x = \pm 1/100$, $f(x) - 1 = 1/10\,000$

ş.a.m.d. Mai general, dacă ε este un număr pozitiv oarecare, oricît de mic, atunci diferenţa dintre $f(x)$ şi 1 va fi mai mică decît ε, cu condiţia ca distanţa dintre x şi 0 să fie mai mică decît numărul $\delta = \sqrt{\varepsilon}$. Într-adevăr dacă

$$|x| < \sqrt{\varepsilon},$$

atunci

$$|f(x) - 1| = |x^2| < \varepsilon.$$

Analogia cu definiţia limitei unui şir este deplină. La p.308 am dat definiţia: *şirul* a_n *are limita* a, *cînd* n *tinde spre infinit, dacă oricărui număr pozitiv* ε, *oricît de mic, îi corespunde un întreg* N (*care depinde de* ε), *astfel încît*

$$|a_n - a| < \varepsilon$$

pentru orice n, *care satisface inegalitatea*

$$n \geqslant N.$$

În cazul unei funcţii $f(x)$ de o variabilă continuă x, atunci cînd x tinde spre o valoare finită x_1, înlocuim doar pe n „suficient de mare", dat de N, prin „*suficient de aproape*" de x_1, dat de un număr δ, şi ajungem la următoarea definiţie a limitei prin apropiere continuă, dată pentru prima dată de Cauchy, prin 1820: *Funcţia* f(x) *are limita* a, *cînd* x *tinde spre valoarea* x_1, *dacă oricărui număr pozitiv* ε, *oricît de mic, îi corespunde un număr pozitiv* δ (*care depinde de* ε), *astfel încît*

$$|f(x) - a| < \varepsilon$$

pentru orice $x \neq x_1$, *care satisface inegalitatea*

$$|x - x_1| < \delta.$$

Dacă se întîmplă acest lucru, scriem

$$f(x) \to a, \text{ cînd } x \to x_1.$$

În cazul funcţiei $f(x) = (x + x^3)/x$, am arătat mai sus că $f(x)$ are limita 1 cînd x tinde spre valoarea $x_1 = 0$. În acest caz a fost suficient să alegem $\delta = \sqrt{\varepsilon}$.

2. Observații asupra noțiunii de limită

Definiția cu (ε, δ) a limitei este rezultatul a mai mult de o sută de ani de încercări și erori, și întruchipează în cîteva cuvinte rezultatul unor eforturi perseverente, făcute pentru a pune această noțiune pe o bază matematică solidă. Noțiunile fundamentale ale analizei — derivata și integrala — pot fi definite numai prin procese de trecere la limită. Dar o înțelegere clară și o definiție precisă a limitelor au fost împiedicate multă vreme de o dificultate aparent de neînlăturat.

În studiul lor asupra mișcării în particular și al oricărei schimbări în general, matematicienii secolelor al XVII-lea și al XVIII-lea au acceptat ca lucru de la sine înțeles noțiunea de cantitate x, care se află în schimbare continuă și care se apropie continuu spre valoarea limită x_1. Legată de această scurgere primară a timpului, sau a unei cantități x, care se comportă ca timpul, ei considerau o valoare secundară $u = f(x)$ care urma mișcarea lui x. Problema consta în a asocia un înțeles matematic precis ideii că $f(x)$ „tinde spre" sau „se apropie de" o valoare fixată a, atunci cînd x se mișcă spre x_1.

Încă de pe vremea lui Zenon și a paradoxurilor sale, noțiunea intuitivă fizică sau metafizică de mișcare continuă a scăpat oricăror încercări de formulare matematică exactă. Nu este nici o dificultate dacă procedăm pas cu pas, trecînd printr-un șir discret de valori a_1, a_2, a_3, ... Dar cînd este vorba de o variabilă continuă x, care parcurge un interval al axei numerice, este imposibil să spunem cum trebuie „să se apropie" x de o valoare fixată x_1, astfel încît el să ia consecutiv, și în ordinea mărimii, toate valorile din acest interval. Într-adevăr, punctele de pe o dreaptă formează o mulțime densă și nu există un punct „următor" după un anumit punct. Desigur, ideea intuitivă de continuu are o realitate psihologică în intelectul uman. Însă ea nu poate fi invocată pentru a rezolva o imposibilitate matematică; inevitabil, rămîne o deosebire între ideea intuitivă și limbajul matematic conceput pentru descrierea faptelor științifice semnificative ale intuiției noastre, în termeni logici preciși. Paradoxurile lui Zenon sugerează această deosebire.

Meritul lui Cauchy a fost tocmai faptul că a înțeles că în ceea ce privește noțiunile matematice, orice referire la ideea intuitivă apriorică de mișcare continuă, poate, și chiar trebuie să fie omisă. Așa cum se întîmplă atît de des, calea spre progresul științific a fost deschisă atunci cînd s-a renunțat la încercarea de a da explicații metafizice, operîndu-se în schimb numai cu noțiuni care, în principiu, corespund unor fenomene „observabile". Dacă analizăm ceea ce înțelegem într-adevăr prin cuvintele „apropiere continuă", cum trebuie să procedăm pentru a o verifica într-un anumit caz, atunci sîntem siliți să acceptăm o definiție ca cea a lui Cauchy. Această definiție este *statică*: ea nu presupune ideea intuitivă de mișcare. Dimpotrivă, doar o astfel de definiție statică face posibilă o analiză matematică precisă a mișcării continue în timp și elimină paradoxurile lui Zenon, în măsura în care ele se referă la știința matematică.

În definiția cu ε și δ, variabila independentă nu se mișcă; ea nu „tinde spre" sau „se apropie de" o limită x_1 în sens fizic. Însă aceste fraze ca și simbolul $\rightarrow$ rămîn un auxiliar intuitiv, și nici un matematician nu trebuie sau nu ar trebui să piardă înțelesul intuitiv sugestiv al faptelor pe care le exprimă. Dar atunci cînd este vorba de a verifica existența unei limite printr-un procedeu științific efectiv, trebuie aplicată definiția cu ajutorul lui ε și δ. Dacă această definiție corespunde în mod satisfăcător noțiunii „dinamice" intuitive de apropiere, aceasta este o problemă de aceeași natură cu problema dacă axiomele geometriei realizează o descriere satisfăcătoare a noțiunii intuitive de spațiu. Ambele formulări neglijează ceva din reprezentările pe care ni le oferă intuiția, dar în schimb ele creează un cadru matematic adecvat pentru a exprima cunoștințele noastre despre aceste noțiuni.

Ca și în cazul limitei unui șir, cheia definiției lui Cauchy se află în inversarea ordinii „naturale" în care sînt considerate variabilele. În primul rînd ne fixăm atenția asupra unei margini ε pentru variabila dependentă și apoi încercăm să determinăm o margine potrivită δ pentru variabila independentă. Afirmația „$f(x) \rightarrow a$ cînd $x \rightarrow x_1$" este doar o exprimare prescurtată a faptului că acest lucru poate fi făcut pentru orice număr pozitiv ε. În particular, nici o *parte* a acestei afirmații, de pildă „$x \rightarrow x_1$", nu are un înțeles în sine.

Mai trebuie subliniat un lucru. Făcînd pe x „să tindă spre" x_1, putem permite ca x să fie mai mare sau mai mic decît x_1, însă excludem în mod intenționat egalitatea, cerînd ca $x \neq x_1$: x tinde spre x_1, dar nu *ia* niciodată valoarea x_1. Astfel, putem aplica definiția dată unor funcții care nu sînt definite pentru $x = x_1$, dar au limite precise, cînd x tinde spre x_1; un exemplu este funcția $f(x) = \dfrac{x + x^3}{x}$, considerată la p. 321. Excluzînd pe $x = x_1$, stabilim o corespondență cu faptul că pentru limitele șirurilor a_n, cînd $n \rightarrow \infty$, de exemplu $a_n = 1/n$, nu facem niciodată substituția $n = \infty$ direct în formulă.

Totuși, cînd x tinde spre x_1, $f(x)$ se poate apropia de limita a, astfel încît există valori $x \neq x_1$, pentru care $f(x) = a$. De exemplu, considerînd funcția $f(x) = x/x$ cînd x tinde către 0, nu permitem niciodată lui x să fie egal cu 0, însă $f(x) = 1$ pentru orice $x \neq 0$ și limita a există și este egală cu 1, în concordanță cu definiția dată.

3. Limita lui $\dfrac{\sin x}{x}$

Dacă x este măsura în radiani a unui unghi, atunci expresia $\dfrac{\sin x}{x}$ este definită pentru orice x, cu excepția lui $x = 0$, în care ea devine simbolul lipsit de sens $0/0$. Cititorul care dispune de o tabelă de funcții trigonometrice

va putea calcula valoarea lui $\frac{\sin x}{x}$ pentru valori mici ale lui x. Aceste tabele sînt date de obicei în funcție de măsura în grade a unghiurilor; reamintim din § 1, secțiunea 2, că măsura în grade x este legată de măsura în radiani y prin relația $x = \frac{\pi}{180} y = 0{,}01745\, y$, cu cinci zecimale exacte. Dintr-o tabelă cu patru zecimale exacte găsim că pentru un unghi de

$10°$	$x = 0{,}1745$,	$\sin x = 0{,}1736$,	$\frac{\sin x}{x} = 0{,}9948$
$5°$,	$0{,}0873$,	$0{,}0872$,	$0{,}9988$
$2°$,	$0{,}0349$,	$0{,}0349$,	$1{,}0000$
$1°$,	$0{,}0175$,	$0{,}0175$,	$1{,}0000$.

Cu toate că aceste cifre sînt exacte pînă la a patra zecimală, observăm că

$$\frac{\sin x}{x} \to 1 \text{ cînd } x \to 0. \tag{1}$$

Vom da acum o demonstrație riguroasă a acestei relații.

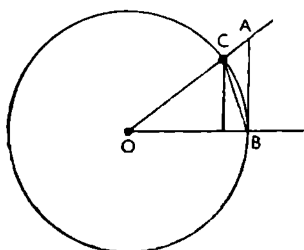


Fig. 169.

Din definiția funcțiilor trigonometrice dată cu ajutorul cercului unitate rezultă că dacă x este măsura în radiani a unghiului BOC, pentru $0 < x < \frac{\pi}{2}$ avem

aria triunghiului $OBC = \frac{1}{2} \cdot 1 \cdot \sin x$,

aria sectorului circular $OBC = \frac{1}{2} \cdot x$ (cf. p. 294),

aria triunghiului $OBA = \frac{1}{2} \cdot 1 \cdot \operatorname{tg} x$.

Deci

$$\sin x < x < \operatorname{tg} x.$$

Împărțind cu $\sin x$, obținem

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x}$$

sau

(2) $$\cos x < \frac{\sin x}{x} < 1.$$

Dar, $1 - \cos x = (1 - \cos x)\dfrac{1 + \cos x}{1 + \cos x} = \dfrac{1 - \cos^2 x}{1 + \cos x} = \dfrac{\sin^2 x}{1 + \cos x} < \sin^2 x.$

Deoarece $\sin x < x$, de aici rezultă că

(3) $$1 - \cos x < x^2$$

sau

$$1 - x^2 < \cos x.$$

Din (2) obținem inegalitatea finală

(4) $$1 - x^2 < \frac{\sin x}{x} < 1.$$

Cu toate că am presupus că $0 < x < \dfrac{\pi}{2}$, această inegalitate este adevărată și pentru $-\dfrac{\pi}{2} < x < 0$, deoarece $\dfrac{\sin(-x)}{-x} = \dfrac{-\sin x}{-x} = \dfrac{\sin x}{x}$ și $(-x)^2 = x^2$.

Relația (1) rezultă imediat din inegalitatea (4). Într-adevăr, diferența dintre $\dfrac{\sin x}{x}$ și 1 este mai mică decît x^2 și aceasta poate fi făcută mai mică decît orice număr ε, dacă alegem pe $|x| < \delta = \sqrt{\varepsilon}$.

Exerciții: 1) Din inegalitatea (3) deduceți relația $\dfrac{1 - \cos x}{x} \to 0$ cînd $x \to 0$.

Găsiți limitele cînd $x \to 0$ ale următoarelor funcții:

2) $\dfrac{\sin^2 x}{x}$. 3) $\dfrac{\sin x}{x(x-1)}$. 4) $\dfrac{\operatorname{tg} x}{x}$. 5) $\dfrac{\sin ax}{x}$.

6) $\dfrac{\sin ax}{\sin bx}$. 7) $\dfrac{x \sin x}{1 - \cos x}$. 8) $\dfrac{\sin x}{x}$ dacă x este măsurat în grade.

9) $\dfrac{1}{x} - \dfrac{1}{\operatorname{tg} x}$. 10) $\dfrac{1}{\sin x} - \dfrac{1}{\operatorname{tg} x}$.

4. Limite cînd $x \to \infty$

Dacă variabila x este suficient de mare, atunci funcția $f(x) = 1/x$ devine oricît de mică sau „tinde spre 0". De fapt, comportarea acestei funcții cînd x crește, este esențial aceeași cu aceea a șirului $1/n$, cînd n crește. Dăm definiția generală: *funcția* f(x) *are limita* a, *cînd* x *tinde spre infinit, și scriem*

$$f(x) \to a \text{ cînd } x \to \infty,$$

dacă oricărui număr pozitiv ε, *oricît de mic, îi corespunde un număr pozitiv* K *(care depinde de* ε*), astfel încît*

$$|f(x) - a| < \varepsilon$$

cu condiția ca $|x| > K$. (Compară cu definiția corespunzătoare de la p.322).

În cazul funcției $f(x) = 1/x$, pentru care $a = 0$, este suficient să alegem $K = 1/\varepsilon$, după cum cititorul poate verifica cu ușurință.

Exerciții: 1) Arătați că definiția precedentă a afirmației

$$f(x) \to a \text{ cînd } x \to \infty$$

este echivalentă cu afirmația $f(x) \to a$ cînd $1/x \to 0$.

Demonstrați că au loc următoarele relații:

2) $\dfrac{x+1}{x-1} \to 1$ cînd $x \to \infty$. 3) $\dfrac{x^2+x+1}{x^2-x-1} \to 1$ cînd $x \to \infty$.

4) $\dfrac{\sin x}{x} \to 0$ cînd $x \to \infty$. 5) $\dfrac{x+1}{x^2+1} \to 0$ cînd $x \to \infty$.

6) $\dfrac{\sin x}{x + \cos x} \to 0$ cînd $x \to \infty$. 7) $\dfrac{\sin x}{\cos x}$ nu are limită cînd $x \to \infty$.

8) Definiți „$f(x) \to \infty$ cînd $x \to \infty$". Dați un exemplu.

Există o deosebire între cazul unei funcții $f(x)$ și al unui șir a_n. În cazul unui șir, n poate tinde spre infinit doar crescînd, însă pentru o funcție putem permite ca x să devină infinit, fie prin valori pozitive, fie prin valori negative. Dacă dorim să ne fixăm atenția asupra comportării lui $f(x)$, cînd x ia numai valori mari *pozitive*, putem înlocui condiția $|x| > K$ prin condiția $x > K$; pentru valori mari negative ale lui x, putem folosi condiția $x < -K$. Pentru a simboliza aceste două metode de apropiere „unilaterală" spre infinit, scriem respectiv

$$x \to +\infty, \quad x \to -\infty.$$

§ 4. DEFINIȚIA PRECISĂ A CONTINUITĂȚII

În § 1, secțiunea 5, am enunțat de fapt următorul criteriu de continuitate a unei funcții: „o funcție $f(x)$ este continuă în punctul $x = x_1$, dacă atunci cînd x se apropie de x_1, cantitatea $f(x)$ se apropie de limita $f(x_1)$". Dacă analizăm această definiție, vedem că ea constă din două condiții diferite.

a) limita a a lui $f(x)$ trebuie să existe cînd x tinde spre x_1.

b) această limită a trebuie să fie egală cu valoarea $f(x_1)$. Dacă în definiția limitei de la p.322 punem $a = f(x_1)$, atunci condiția continuității capătă următoarea formă: *funcția* f(x) *este continuă pentru valoarea* $x = x_1$, *dacă oricărui număr pozitiv* ε *oricît de mic, îi corespunde un număr pozitiv* δ (*care depinde de* ε), *astfel încît*

$$|f(x) - f(x_1)| < \varepsilon$$

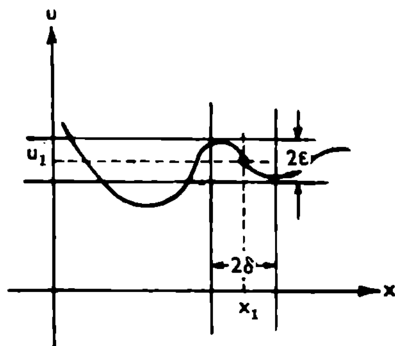


Fig. 170. Funcție continuă în $x = x_1$

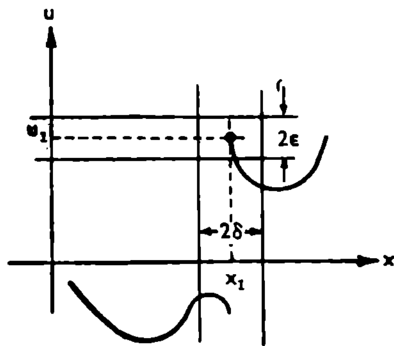


Fig. 171. Funcție discontinuă în $x = x_1$

pentru orice x *care satisface inegalitatea*

$$|x - x_1| < \delta.$$

(Condiția $x \neq x_1$ impusă în definiția limitei nu este necesară aici, pentru că inegalitatea $|f(x_1) - f(x_1)| < \varepsilon$ este îndeplinită întotdeauna.)

Ca exemplu, să verificăm continuitatea funcției $f(x) = x^3$ în punctul $x_1 = 0$. Avem

$$f(x_1) = 0^3 = 0.$$

Să dăm acum lui ε o valoare pozitivă oarecare, de exemplu $\varepsilon = \frac{1}{1000}$. Trebuie să arătăm că restrîngînd pe x la valori suficient de apropiate de $x_1 = 0$, valorile corespunzătoare ale lui $f(x)$ nu vor diferi de 0 cu mai mult de $\frac{1}{1000}$, adică se vor afla între $-\frac{1}{1000}$ și $+\frac{1}{1000}$. Se vede imediat că această margine nu este depășită dacă restrîngem pe x la valori care diferă de $x_1 = 0$ cu mai

puțin de $\delta = \sqrt[3]{\frac{1}{1000}} = \frac{1}{10}$, deoarece dacă $|x| < \frac{1}{10}$, atunci $|f(x)| = |x|^3 < \frac{1}{1000}$. În același mod putem înlocui pe $\varepsilon = \frac{1}{1000}$ prin $\varepsilon = 10^{-4}, 10^{-5}$ sau oricare altă margine dorim; $\delta = \sqrt[3]{\varepsilon}$ va satisface întotdeauna condiția dată, deoarece dacă $|x| < \sqrt[3]{\varepsilon}$, atunci $|f(x)| = |x|^3 < \varepsilon$.

Pe baza definiției continuității cu ajutorul lui ε și δ, putem arăta în mod analog că toate polinoamele, funcțiile raționale și funcțiile trigonometrice sînt continue, cu excepția unor valori izolate ale lui x, în care funcțiile pot deveni infinite.

Cu ajutorul graficului unei funcții $u = f(x)$, definiția continuității capătă următoarea formă geometrică. Alegeți un număr pozitiv ε și trasați paralele la axa Ox, la distanțe egale cu $f(x_1) - \varepsilon$ și $f(x_1) + \varepsilon$. Atunci trebuie să fie posibilă găsirea unui număr pozitiv δ, astfel încît întreaga porțiune a graficului, care se află în interiorul benzii verticale de lățime 2δ în jurul lui x_1 să fie conținută în interiorul benzii orizontale de lățime 2ε, în jurul lui $f(x_1)$. Fig. 170 arată o funcție care este continuă în x_1, în timp ce fig. 171 arată o funcție care nu este continuă. În ultimul caz, oricît de îngustă am face banda verticală din jurul lui x_1, ea va include întotdeauna o porțiune a graficului care se află în exteriorul benzii orizontale, corespunzătoare unei alegeri potrivite a lui ε.

Dacă afirm că o funcție $u = f(x)$ este continuă pentru valoarea $x = x_1$, aceasta înseamnă că sînt gata să aduc la îndeplinire următorul contract făcut cu dv. Puteți alege orice număr pozitiv ε, oricît de mic doriți, dar fixat. Atunci eu trebuie să găsesc un număr pozitiv δ, astfel încît dacă $|x - x_1| < \delta$, atunci $|f(x) - f(x_1)| < \varepsilon$. Nu mă angajez să găsesc de la început un număr δ care va satisface condiția pentru orice ε pe care l-ați putea alege după aceea; modul în care *eu* aleg pe δ va depinde de modul în care *dv.* alegeți pe ε. Dacă dv. puteți găsi o valoare a lui ε, pentru care eu nu pot indica un δ potrivit, atunci afirmația mea este contrazisă De acee a, pentru a demonstra că îmi pot îndeplini contractul în orice caz concret al une funcții $u=f(x)$, de obicei construiesc o funcție pozitivă explicită

$$\delta = \varphi(\varepsilon),$$

definită pentru orice număr pozitiv ε, pentru care pot arăta că $|x - x_1| < \delta$ implică întotdeauna $|f(x) - f(x_1)| < \varepsilon$. În cazul funcției $u = f(x) = x^3$, în $x_1 = 0$, funcția $\delta = \varphi(\varepsilon)$ era $\delta = \sqrt[3]{\varepsilon}$.

Exerciții: 1) Arătați că $\sin x$, $\cos x$ sînt funcții continue.

2) Demonstrați continuitatea funcțiilor $1/(1 + x^4)$ și $\sqrt{1 + x^2}$.

Acum este clar că definiția cu ajutorul lui ε și δ a continuității concordă cu ceea ce s-ar putea numi faptele observabile referitoare la o funcție. Ca atare, ea este în concordanță cu principiul general al științei moderne, care pune drept criteriu al utilității unei noțiuni sau al „existenței științifice" a unui fenomen posibilitatea observării lui (cel puțin în principiu) sau a reducerii lui la fapte observabile.

§ 5. DOUĂ TEOREME FUNDAMENTALE REFERITOARE LA FUNCŢII CONTINUE

1. Teorema lui Bolzano

Bernard Bolzano (1781—1848), un preot catolic, cunoscător al filozofiei scolastice, a fost unul dintre primii care au introdus în analiza matematică noţiunea modernă de rigoare. Importanta sa carte *Paradoxien des Unendlichen* a apărut în 1850. Aici s-a recunoscut pentru prima dată că multe afirmaţii aparent evidente, referitoare la funcţiile continue, pot şi trebuie să fie demonstrate, dacă vrem ca ele să fie folosite în toată generalitatea lor. Teorema următoare, referitoare la funcţii continue de o variabilă, constituie un exemplu.

O funcţie continuă de o variabilă x, *care este pozitivă pentru o valoare a lui* x *şi negativă pentru altă valoare a lui* x *din intervalul închis* $a \leqslant x \leqslant b$ *de continuitate a funcţiei, trebuie să se anuleze pentru o anumită valoare intermediară a lui* x. Astfel, dacă $f(x)$ este continuă cînd x variază de la a la b şi dacă $f(a) < 0$ şi $f(b) > 0$, atunci va exista o valoare α a lui x, astfel încît $a < \alpha < b$ şi $f(\alpha) = 0$.

Teorema lui Bolzano corespunde perfect ideii noastre intuitive de curbă continuă, care pentru a trece de la un punct aflat sub axa Ox la un punct aflat deasupra ei, trebuie să traverseze axa cel puţin într-un punct. Faptul că acest lucru nu este neapărat adevărat pentru funcţiile discontinue, este arătat de fig. 157 de la p. 301.

*2. Demonstraţia teoremei lui Bolzano

În cele ce urmează va fi dată o demonstraţie riguroasă a acestei teoreme. (Ca şi Gauss şi ca alţi mari matematicieni, o putem accepta şi folosi fără demonstraţie.) Scopul nostru este de a reduce teorema la proprietăţile fundamentale ale sistemului numerelor reale, în particular la postulatul lui Dedekind-Cantor, referitor la şirurile descrescătoare de intervale (*p.* 84). În acest scop, considerăm intervalul I, $a \leqslant x \leqslant b$, în care este definită funcţia $f(x)$ şi îl înjumătăţim, determinînd punctul $x_1 = \frac{a+b}{2}$. Dacă în acest punct găsim că $f(x_1) = 0$, atunci nu mai rămîne nimic de arătat. Dacă însă $f(x_1) \neq 0$, atunci $f(x_1)$ trebuie să fie sau mai mare, sau mai mic decît zero. În ambele cazuri, una dintre jumătăţile lui I va avea şi ea proprietatea că semnul lui $f(x)$ este diferit la cele două extremităţi. Să notăm cu I_1 acest interval. Continuăm procesul, înjumătăţind pe I_1; atunci, sau $f(x) = 0$ în mijlocul lui I_1, sau putem alege un interval I_2, o jumătate a lui I_1, cu proprietatea că semnul lui $f(x)$ este diferit la cele două extremităţi. Repetînd acest procedeu, sau vom găsi după un număr finit de înjumătăţiri, un punct pentru care $f(x) = 0$, sau vom obţine un şir descrescător de intervale $I_1, I_2, I_3, \ldots$. În ultimul caz, postulatul lui Dedekind-Cantor asigură existenţa unui punct α, comun tuturor acestor intervale. Afirmăm că $f(\alpha) = 0$, astfel încît α este punctul a cărui existenţă demonstrează teorema.

Pînă acum nu a fost utilizată ipoteza continuităţii. Ea va fi folosită acum pentru a încheia demonstraţia printr-un raţionament indirect. Vom demonstra că $f(\alpha) = 0$, presupunînd contrariul şi deducînd o contradicţie. Să presupunem că $f(\alpha) \neq 0$, de pildă $f(\alpha) = 2\varepsilon > 0$. Deoarece $f(x)$ este continuă, putem găsi un interval J, de lungime 2δ (eventual foarte mică), cu centrul în α, astfel încît valoarea lui $f(x)$ în orice punct din J să difere de $f(\alpha)$ cu mai puţin decît ε. Deci

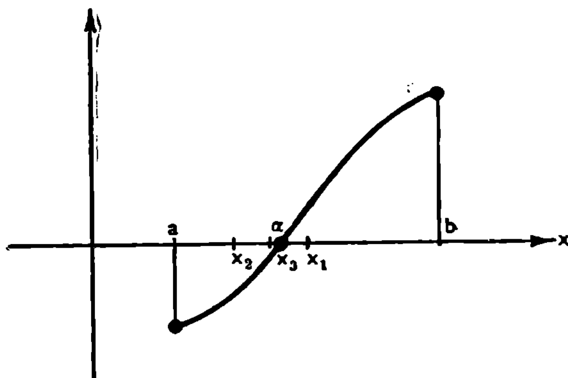


Fig. 172. Teorema lui Bolzano

deoarece $f(\alpha) = 2\varepsilon$, putem fi siguri că $f(x) > \varepsilon$ în orice punct din J, astfel încît $f(x) > 0$ în J. Însă intervalul J este fixat şi, dacă n este suficient de mare, intervalul I_n trebuie să se afle în mod necesar în interiorul lui J, dat fiind că şirul I_n se contractă spre zero. Aceasta duce la contradicţie, pentru că din modul în care a fost ales I_n rezultă că funcţia $f(x)$ are valori de semn opus la cele două extremităţi ale fiecărui I_n, astfel încît $f(x)$ trebuie să aibă valori negative în cel puţin un punct din J. Deci absurditatea ipotezei $f(\alpha) > 0$ şi a ipotezei $f(\alpha) < 0$ (care se demonstrează în acelaşi mod) arată că $f(\alpha) = 0$.

3. Teorema lui Weierstrass asupra valorilor extreme

Un alt fapt important şi plauzibil din punct de vedere intuitiv, referitor la funcţiile continue, a fost formulat de Karl Weierstrass (1815—1897) care, poate mai mult decît oricare altul, are meritul tendinţei moderne spre rigoare în analiza matematică. Această teoremă afirmă: *dacă o funcţie* f(x) *este continuă într-un interval* I, $a \leqslant x \leqslant b$, *inclusiv la extremităţile* a *şi* b *ale intervalului, atunci există cel puţin un punct în* I, *în care* f(x) *îşi atinge cea mai mare valoare* M *şi un alt punct, în care* f(x) *îşi atinge cea mai mică valoare* m. Intuitiv vorbind, aceasta înseamnă că graficul funcţiei continue $u = f(x)$ trebuie să aibă cel puţin un punct de înălţime maximă şi un altul de înălţime minimă.

Este important să observăm că afirmaţia nu mai este neapărat adevărată, dacă funcţia $f(x)$ încetează de a mai fi continuă în extremităţile lui I. De

exemplu, funcția $f(x) = 1/x$ nu are o cea mai mare valoare în intervalul $0 < x \leqslant 1$, cu toate că $f(x)$ este continuă în interiorul acestui interval. Nici funcția discontinuă nu ia neapărat o cea mai mare sau o cea mai mică valoare, chiar dacă este mărginită. De exemplu, să considerăm funcție discontinuă $f(x)$, definită prin

$$f(x) = x \text{ pentru } x \text{ irațional,}$$

$$f(x) = \frac{1}{2} \text{ pentru } x \text{ rațional,}$$

în intervalul $0 \leqslant x \leqslant 1$. Această funcție ia întotdeauna valori cuprinse între 0 și 1 și de fapt ia valori oricît de apropiate vrem de 0 și de 1, dacă x este ales ca număr irațional, suficient de aproape de 0 sau de 1. Însă $f(x)$ nu poate fi niciodată *egală* cu 0 sau cu 1, deoarece pentru valori raționale ale lui x avem $f(x) = \frac{1}{2}$, iar pentru valori iraționale ale lui x avem $f(x) = x$. Deci 0 și 1 nu sînt atinse niciodată.

*Teorema lui Weierstrass poate fi demonstrată într-un mod analog cu cel al teoremei lui Bolzano. Împărțim intervalul I în două jumătăți închise I' și I'' și ne fixăm atenția asupra lui I' ca interval în care trebuie să căutăm cea mai mare valoare a lui $f(x)$, *afară de cazul în care există un punct* α *în* I'', *astfel încît* $f(\alpha)$ *să depășească toate valorile pe care* $f(x)$ *le ia în* I'; în acest caz îl alegem pe I''. Să notăm cu I_1 intervalul ales în acest mod. Acum procedăm cu I_1 tot așa cum am procedat cu I, obținînd un interval I_2 ș.a.m.d. Acest proces va defini un șir $I_1, I_2, \ldots .I_n, \ldots$ de intervale, care conțin toate un punct z. Vom demonstra că valoarea $f(x) = M$ este cea mai mare dintre valorile lui $f(x)$ în I, adică nu poate exista nici un punct s din I pentru care $f(x) > M$. Să presupunem că există un punct s în care $f(s) = M + 2\varepsilon$, unde ε este număr pozitiv (eventual foarte mic). Cu centrul în z, datorită continuității lui $f(x)$. putem alege un interval K care nu îl conține pe s, astfel încît valorile lui $f(x)$ luate în K să difere de $f(z) = M$ cu mai puțin decît ε; deci vom avea cu siguranță $f(x) < M + \varepsilon$ în orice punct din K. Dar pentru un n suficient de mare, intervalul I_n se află în interiorul lui K și I_n a fost definit astfel, încît nici o valoare a lui $f(x)$, luată în exteriorul lui I_n, nu poate întrece toate valorile pe care $f(x)$ le ia în I_n. Deoarece s este exterior lui I_n și $f(s) > M + \varepsilon$, în timp ce în K, și deci și în I_n avem $f(x) < M + \varepsilon$, am ajuns la o contradicție.

Existența unei cele mai mici valori m poate fi demonstrată în același mod sau ea rezultă imediat din ceea ce s-a demonstrat, deoarece cea mai mică valoare a lui $f(x)$ este cea mai mare valoare a lui $g(x) = -f(x)$.

Teorema lui Weierstrass poate fi demonstrată într-un mod asemănător pentru funcții continue de două sau mai multe variabile $x, y, \ldots$. În locul unui interval împreună cu extremitățile trebuie să considerăm un domeniu *închis*, ca de pildă un dreptunghi din planul x, y, împreună cu frontiera lui.

Exercițiu: Unde am folosit în demonstrațiile teoremelor lui Bolzano și Weierstrass faptul că $f(x)$ a fost presupusă definită și continuă, în întreg intervalul *închis* $a \leqslant x \leqslant b$, și nu numai în intervalul $a < x \leqslant b$ sau $a < x < b$?

Demonstraţiile teoremelor lui Bolzano şi Weierstrass au o natură neconstructivă. Ele nu indică o metodă pentru găsirea efectivă a unui zero, sau a celei mai mari, sau a celei mai mici valori a unei funcţii cu un grad prescris de precizie, printr-un număr finit de paşi. Se demonstrează doar existenţa sau mai degrabă absurditatea neexistenţei valorilor dorite. Aici este un alt exemplu important în care „intuiţioniştii" (cf. p. 103) au adus obiecţii; unii au insistat chiar ca astfel de teoreme să fie eliminate din matematică. Cel ce studiază matematica n-ar trebui să privească aceste obiecţii cu mai multă seriozitate decît au făcut-o cei mai mulţi dintre critici.

*4. O teoremă referitoare la şiruri. Mulţimi compacte

Fie $x_1, x_2, x_3, \ldots$ un şir infinit de numere, distincte sau nu, conţinute în *intervalul închis* I, $a \leqslant x \leqslant b$. Şirul poate să tindă sau nu către o limită. Însă în orice caz, *întotdeauna este posibil să extragem dintr-un astfel de şir, omiţînd o parte din termenii lui, un subşir infinit* $y_1, y_2, y_3, \ldots$, *care tinde spre o limită* y, *conţinută în intervalul* I.

Pentru a demonstra această teoremă, împărţim intervalul I în două subintervale închise I' şi I'', cu ajutorul mijlocului $\frac{a+b}{2}$ al lui I:

$$I' : a \leqslant x \leqslant \frac{a+b}{2},$$

$$I'' : \frac{a+b}{2} \leqslant x \leqslant b.$$

Cel puţin într-unul dintre ele, pe care îl putem nota cu I_1, trebuie să se afle o infinitate de termeni x_n din şirul iniţial. Să alegem unul dintre aceşti termeni, de pildă pe x_{n_1} şi să-l notăm cu y_1. Să procedăm în acelaşi mod cu intervalul I_1. Deoarece există o infinitate de termeni x_n în I_1, trebuie să existe o infinitate de termeni în cel puţin una dintre jumătăţile lui I_1, pe care o putem nota cu I_2. Deci putem găsi un termen x_n în I_2, pentru care $n > n_1$. Să alegem unul dintre aceştia şi să-l notăm cu y_2. Procedînd în acest mod, putem găsi un şir descrescător $I_1, I_2, I_3, \ldots$ de intervale şi un subşir $y_1, y_2, y_3, \ldots$ al şirului iniţial, astfel încît y_n să se afle în I_n, pentru orice n. Acest şir de intervale se contractă spre un punct y din I şi este limpede că şirul $y_1, y_2, y_3, \ldots$ are limita y, ceea ce trebuia demonstrat.

*Aceste raţionamente pot fi generalizate într-un mod tipic matematicii moderne. Să considerăm o variabilă x, care parcurge o mulţime S, în care este definită o noţiune de „distanţă". S poate fi o mulţime de puncte din plan sau din spaţiu, dar acest lucru nu este necesar;

de exemplu, S poate fi mulțimea tuturor triunghiurilor din plan. Dacă X și Y sînt două triunghiuri, cu vîrfurile respectiv în A, B, C și A', B', C', atunci putem defini „distanța" dintre două triunghiuri, ca fiind numărul

$$d(X, Y) = AA' + BB' + CC',$$

unde AA' etc. este distanța obișnuită dintre punctele A și A'. Ori de cîte ori există o astfel de noțiune de „distanță" într-o mulțime S, putem defini noțiunea de șir de elemente X_1, X_2, X_3, ... care tinde spre un element limită X din S. Prin aceasta înțelegem că $d(X, X_n) \to 0$ cînd $n \to \infty$. Vom spune acum că *mulțimea* S *este compactă dacă din orice șir* $X_1, X_2, X_3, \ldots$ *de elemente din* S *putem extrage întotdeauna un subșir, care tinde spre un element limită* X *din* S. În cele de mai sus am arătat că un interval închis $a \leqslant x \leqslant b$ este compact în acest sens. Prin urmare, noțiunea de mulțime compactă poate fi privită ca generalizare a *intervalului închis* de pe axa numerică. Remarcați că axa numerică *în întregime* nu este compactă, deoarece șirul de întregi 1, 2, 3, 4, 5, ... nici nu tinde către o limită și nici nu conține vreun subșir convergent. Nici un interval deschis, ca de pildă $0 < x < 1$, care nu își conține extremitățile, nu este compact, deoarece șirul $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ sau orice subșir al său tinde spre limita 0, care nu este un punct al intervalului deschis. În același mod, se poate arăta că porțiunea din plan formată din punctele interioare unui pătrat sau unui dreptunghi nu este compactă, dar devine compactă dacă îi adăugăm punctele de pe frontieră. Mai mult, mulțimea tuturor triunghiurilor, ale căror vîrfuri se află în interiorul sau pe circumferința unui cerc dat, este compactă.

Putem extinde, de asemenea, noțiunea de continuitate la cazul în care variabila X parcurge orice mulțime S, în care este definită noțiunea de limită. Funcția $u = F(X)$, unde u este un număr real, se numește continuă în elementul X, dacă pentru orice șir de elemente $X_1, X_2, X_3, \ldots$ care tinde spre limita X, șirul corespunzător de numere $F(X_1)$, $F(X_2)$ ... tinde spre limita $F(X)$. (Se mai poate da și o definiție echivalentă, cu ajutorul lui ε și δ.) Este foarte ușor de arătat că teorema lui Weierstrass rămîne în vigoare și în cazul general al unei funcții continue, definite pe elementele unei mulțimi compacte:

Dacă u = F(X) *este o funcție continuă, definită pe o mulțime compactă* S, *atunci există întotdeauna un element al lui* S *pentru care* F(X) *își atinge cea mai mare valoare și unul în care își atinge cea mai mică valoare.*

Demonstrația nu prezintă nici un fel de dificultăți, deoarece am luat cunoștință de noțiunile generale care intervin; nu vom continua însă cu acest subiect. În cap. VII vom vedea că teorema generală a lui Weierstrass are o mare importanță în teoria maximelor și minimelor.

§ 6. CÎTEVA APLICAȚII ALE TEOREMEI LUI BOLZANO

1. Aplicații geometrice

Teorema simplă, dar generală, a lui Bolzano poate fi folosită pentru a demonstra multe lucruri care nu sînt cîtuși de puțin evidente la prima vedere. Începem demonstrînd că: *dacă* A *și* B *sînt două suprafețe din plan, atunci exis-*

tă o dreaptă în plan, care înjumătăţeşte simultan ariile lui A *şi* B. Prin „suprafaţă" înţelegem orice porţiune a planului, inclusă în interiorul unei curbe simple închise.

Să începem alegînd un punct fix P în plan şi trasînd din P o semidreaptă PR, pe care o vom lua ca origine pentru măsurarea unghiurilor. Dacă luăm o semidreaptă PS care face un unghi x cu PR, atunci va exista o dreaptă în plan

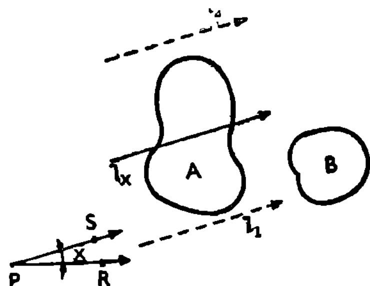


Fig. 173. Înjumătăţirea simultană a două suprafeţe

care *înjumătăţeşte aria lui* A, şi care are aceeaşi direcţie ca şi semidreapta PS. Într-adevăr, să luăm o dreaptă l_1, de aceeaşi direcţie cu PS şi care se află în întregime de-o parte a lui A şi să deplasăm această dreaptă paralel cu ea însăşi pînă ce ajunge în poziţia l_2 (fig. 173), aflată în întregime de cealaltă parte a lui A. Atunci funcţia, a cărei valoare este egală cu aria lui A aflată la dreapta dreptei mobile (orientată în acelaşi sens cu semidreapta dusă din P), din care se scade aria lui A aflată la stînga dreptei, va fi pozitivă pentru poziţia l_1 şi negativă pentru poziţia l_2. Deoarece această funcţie este continuă, din teorema lui Bolzano rezultă că ea trebuie să se anuleze pentru o poziţie intermediară l_x, care prin urmare înjumătăţeşte suprafaţa A. Pentru orice valoare a lui x, cuprinsă între $x = 0°$ şi $x = 360°$, dreapta l_x, care înjumătăţeşte suprafaţa A, este determinată în mod unic.

Acum să definim funcţia $y = f(x)$, ca fiind aria lui B, aflată la dreapta lui l_x, din care se scade aria lui B aflată la stînga lui l_x. Să presupunem că dreapta l_0, care înjumătăţeşte suprafaţa A şi are direcţia lui PR, lasă o porţiune mai mare a lui B la dreapta ei decît la stînga; atunci, pentru $x = 0°$, y este pozitiv. Să facem pe x să crească pînă la 180°; atunci dreapta l_{180} cu direcţia RP, care înjumătăţeşte pe A, este aceeaşi cu dreapta l_0, dar are direcţie opusă, dreapta şi stînga fiind permutate. Deci valoarea numerică a lui y pentru $x = 180°$ este aceeaşi ca şi pentru $x = 0°$, dar de semn opus şi deci negativă. Deoarece y este o funcţie continuă de x, cînd l_x se roteşte, există o valoare α a lui x, cuprinsă între 0 şi 180°, pentru care $y = 0$. Rezultă că dreapta l_α înjumătăţeşte simultan pe A şi pe B. Cu aceasta se încheie demonstraţia.

Să remarcăm că deşi am demonstrat *existenţa* unei drepte cu proprietatea dorită, nu am indicat un anumit procedeu de *construire* a ei; acesta este un alt exemplu al caracterului distinct al demonstraţiei de existenţă matematică faţă de construcţiile efective.

O problemă asemănătoare este următoarea: fiind dată o suprafaţă în plan, se cere să secţionăm aria acestei suprafeţe în *patru* părţi egale, cu ajutorul a

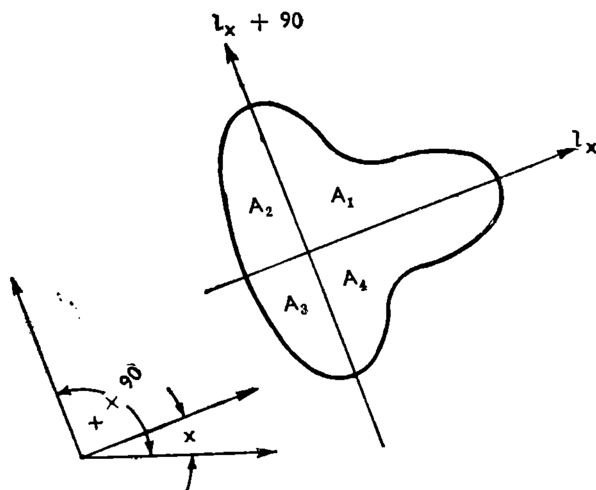


Fig. 174.

două drepte *perpendiculare*. Pentru a arăta că acest lucru este posibil întotdeauna, revenim la problema precedentă la momentul în care am definit pe l_x pentru orice unghi x, dar uităm de existenţa suprafeţei B. În schimb, luăm dreapta l_{x+90}, care este perpendiculară pe l_x şi care înjumătăţeşte şi ea pe A. Dacă numărăm cele patru porţiuni ale lui A, aşa cum se arată în fig. 174, atunci avem

$$A_1 + A_2 = A_3 + A_4$$

şi

$$A_2 + A_3 = A_1 + A_4,$$

de unde rezultă, scăzînd a doua egalitate din prima, că avem

$$A_1 - A_3 = A_3 - A_1,$$

adică

$$A_1 = A_3,$$

şi deci

$$A_2 = A_4.$$

Astfel, dacă putem demonstra existenţa unui unghi α cu proprietatea că pentru l_x avem

$$A_1(\alpha) = A_2(\alpha),$$

atunci teorema va fi demonstrată, deoarece pentru un astfel de unghi cele patru arii vor fi egale. Pentru a face acest lucru, definim o funcţie $y = f(x)$, trasînd pe l_x şi punînd

$$f(x) = A_1(x) - A_2(x).$$

Pentru $x = 0°$, $f(0) = A_1(0) - A_2(0)$ poate fi pozitiv. În acest caz, pentru $x = 90°$, $A_1(90) - A_2(90) = A_2(0) - A_3(0) = A_2(0) - A_1(0)$ va fi negativ. De aceea, deoarece $f(x)$ variază continuu cînd x creşte de la 0 la 90°, va exista o valoare α, cuprinsă între 0 şi 90°, pentru care $f(\alpha) = A_1(\alpha) - A_2(\alpha) = 0$. Atunci dreptele l_α şi $l_{\alpha+90}$ vor împărţi aria suprafeţei în patru părţi egale.

Este interesant de observat că aceste probleme pot fi generalizate la trei sau mai multe dimensiuni. În spaţiul cu trei dimensiuni, prima problemă se formulează astfel: date fiind trei volume în spaţiu, găsiţi un plan care le înjumătăţeşte simultan pe toate trei. Demonstrarea faptului că acest lucru este posibil întotdeauna depinde din nou de teorema lui Bolzano. În spaţii cu mai mult de trei dimensiuni, teorema rămîne adevărată, dar demonstraţia necesită metode mai avansate.

*2. Aplicaţie la o problemă de mecanică

Vom încheia acest paragraf discutînd o problemă aparent dificilă de mecanică, care se rezolvă cu uşurinţă printr-un raţionament bazat pe continuitate. (Această problemă a fost sugerată de H. Whitney.)

Să presupunem că un tren se deplasează de la o staţie A la o staţie B, pe o porţiune rectilinie a căii ferate. Nu este necesar ca mersul trenului să se efectueze cu viteză sau cu acceleraţie constantă. Trenul poate merge în orice mod, accelerînd, încetinind, oprindu-se sau chiar întorcîndu-se înapoi în anumite momente, înainte de a ajunge în B. Se presupune însă că se cunoaşte de la început mişcarea exactă a trenului; adică este dată funcţia $s = f(t)$, unde s este distanţa de la staţia A la tren, iar t este timpul măsurat din momentul plecării. Pe podeaua unuia dintre vagoane este prinsă, printr-o articulaţie, o bară care se poate deplasa fără frecare, înainte sau înapoi, pînă la atingerea podelei. Dacă ea atinge podeaua, presupunem că rămîne nemişcată din acel

moment; acest lucru se va întîmpla dacă bara nu sare. Este oare posibil să așezăm bara într-o astfel de poziție, așa încît dacă ea este lăsată liberă în momentul în care trenul pornește și dacă i se permite să se miște numai sub influența gravitației și a accelerației trenului, să nu cadă pe podea în tot cursul călătoriei de la A spre B?

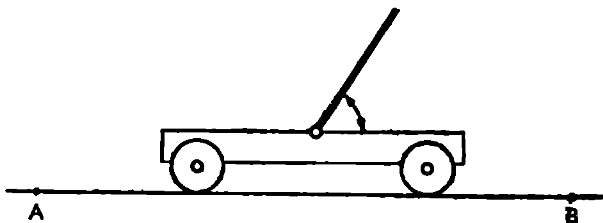


Fig. 175.

Ar putea să pară cu totul improbabil ca pentru orice program dat al deplasării trenului, interacțiunea dintre gravitație și forțele de inerție să permită întotdeauna menținerea unui echilibru, cu singura condiție ca poziția inițială a barei să fie aleasă în mod convenabil. Și totuși afirmăm că o astfel de poziție există întotdeauna.

Oricît de paradoxală ar putea să pară această afirmație la prima vedere, ea poate fi demonstrată cu ușurință îndată ce ne fixăm atenția asupra caracterului ei esențial topologic. Nu este necesară nici o cunoaștere amănunțită a legilor dinamicii, trebuie să admitem doar următoarea ipoteză simplă de natură fizică: *mișcarea barei depinde continuu de poziția ei inițială.* Să caracterizăm poziția inițială a barei prin unghiul inițial x, pe care îl face cu podeaua, și să notăm cu y unghiul pe care bara îl face cu podeaua la sfîrșitul călătoriei, cînd trenul ajunge în punctul B. Dacă bara a căzut pe podea, avem fie $y = 0$, fie $y = \pi$. Pentru o anumită poziție inițială x, poziția finală y, conform ipotezei făcute, este determinată în mod unic ca funcție $y = g(x)$, continuă și care are valorile $y = 0$ pentru $x = 0$ și $y = \pi$ pentru $x = \pi$. (Ultima afirmație exprimă doar faptul că bara rămîne pe podea, dacă se află în această poziție la începutul călătoriei.) Acum ne reamintim faptul că $g(x)$, ca funcție continuă în intervalul $0 \leqslant x \leqslant \pi$, ia toate valorile cuprinse între $g(0) = 0$ și $g(\pi) = \pi$. Prin urmare, pentru orice astfel de valori y, ca de pildă pentru valoarea $y = \frac{\pi}{2}$, există o valoare a lui x, astfel încît $g(x) = y$; în particular, există o poziție inițială pentru care poziția finală a barei în B este perpendiculară pe podea. (Remarcați că în acest raționament nu trebuie să se uite că mișcarea trenului este dată de la început.)

Desigur, raționamentul este cu desăvîrșire teoretic. Dacă călătoria durează multă vreme sau dacă mersul trenului, exprimat prin $s = f(t)$, este foarte

neregulat, atunci domeniul pozițiilor inițiale ale lui x, pentru care poziția finală $g(x)$ diferă de 0 sau de π, va fi foarte mic, după cum își imaginează oricine care a încercat să țină în echilibru un ac vertical pe o farfurie pentru un timp mai îndelungat. Și totuși raționamentul dat ar trebui să prezinte interes și pentru persoanele cu o înclinație practică, deoarece el arată cum se pot obține rezultate calitative în dinamică, cu ajutorul unor raționamente simple, lipsite de complicații tehnice.

Exerciții: 1) Folosind teorema de la p. 333, arătați că raționamentul de mai sus poate fi generalizat, pentru cazul în care călătoria are o durată infinită.

2) Generalizați rezultatul precedent, pentru cazul în care mișcarea trenului se face în lungul unei curbe oarecare din plan, iar bara poate cădea în orice direcție. (Indicație: Nu este posibil să aplicăm continuu un disc circular pe circumferința lui, printr-o aplicație care lasă neschimbat orice punct de pe circumferință (cf. p. 271).

3) Arătați că timpul necesar barei pentru a cădea pe podea, în cazul în care vagonul este nemișcat, iar bara este eliberată dintr-o poziție care face unghiul ε cu verticala, tinde spre infinit, atunci cînd ε tinde spre zero.

ALTE EXEMPLE REFERITOARE LA LIMITE ȘI CONTINUITATE

§ 1. EXEMPLE DE LIMITE

1. Observații generale

În multe cazuri, convergența unui șir a_n poate fi demonstrată printr-un raționament de tipul următor. Găsim alte două șiruri b_n și c_n, ai căror termeni au o structură mai simplă decît cea a termenilor șirului inițial, astfel încît

$$b_n \leqslant a_n \leqslant c_n \tag{1}$$

pentru orice n. *Atunci, dacă putem arăta că șirurile* b_n *și* c_n *converg amîndouă spre aceeași limită* α, *rezultă că și* a_n *converge spre limita* α. Lăsăm pe seama cititorului demonstrarea formală a acestei afirmații.

Este clar că aplicațiile acestui procedeu vor implica folosirea inegalităților. De aceea, este potrivit să reamintim cîteva reguli elementare, care guvernează operațiile aritmetice cu inegalități.

1. Dacă $a > b$, atunci $a + c > b + c$ (orice număr poate fi adăugat ambilor membri ai unei inegalități).

2. Dacă $a > b$ și dacă numărul c este *pozitiv*, atunci $ac > bc$ (o inegalitate poate fi înmulțită cu orice număr pozitiv).

3. Dacă $a < b$, atunci $-b < -a$ (sensul unei inegalități se schimbă dacă înmulțim cu -1 ambii membri ai săi); astfel $2 < 3$, însă $-3 < -2$.

4. Dacă a și b au același semn și dacă $a < b$, atunci $1/a > 1/b$.

5. $|a + b| \leqslant |a| + |b|$.

2. Limita lui q^n

Dacă q este un număr mai mare decît 1, șirul q^n va crește fără margini, așa cum se întîmplă cu șirul $2, 2^2, 2^3, \ldots$ pentru $q = 2$. Șirul „tinde spre infinit" (cf. p. 311). Demonstrarea în cazul general se bazează pe o inegalitate importantă (demonstrată la p. 32):

$$(1 + h)^n \geqslant 1 + nh > nh, \tag{2}$$

unde h este un număr pozitiv oarecare. Punem $q = 1 + h$, unde $h > 0$; atunci

$$q^n = (1 + h)^n > nh.$$

Dacă k este un număr pozitiv oarecare, oricît de mare, atunci pentru orice $n > k/h$ rezultă că

$$q^n > nh > k;$$

deci $q^n \to \infty$.

Dacă $q = 1$, atunci termenii şirului q^n sînt toţi egali cu 1 şi de aceea limita şirului este egală cu 1. Dacă q este negativ, atunci q^n va alterna între valori pozitive şi negative şi nu va avea limită dacă $q \leqslant -1$.

Exerciţiu: Daţi o demonstraţie riguroasă ultimei afirmaţii.

La p. 80 am arătat că dacă $-1 < q < 1$, atunci $q^n \to 0$. Putem da o altă demonstraţie foarte simplă a acestui fapt. Considerăm mai întîi cazul în care $0 < q < 1$. Atunci numerele $q, q^2, q^3, \ldots$ formează un şir descrescător, mărginit inferior de 0. Deci conform celor arătate la p.312, şirul trebuie să tindă spre o limită: $q^n \to a$. Înmulţind ambii membri ai acestei relaţii cu q, obţinem $q^{n+1} \to aq$.

Acum q^{n+1} trebuie să aibă aceeaşi limită ca şi q^n, deoarece numele n sau $n + 1$ al exponentului crescător nu are importanţă. Deci $aq = a$ sau $a(q - 1) = 0$. Deoarece $1 - q \neq 0$, aceasta implică faptul că $a = 0$.

Dacă $q = 0$, afirmaţia $q^n \to 0$ este banală. Dacă $-1 < q < 0$, atunci $0 < |q| < 1$; deci $|q^n| = |q|^n \to 0$ în baza raţionamentului precedent. De aici rezultă că întotdeauna $q^n \to 0$ dacă $|q| < 1$. Cu aceasta demonstraţia s-a încheiat.

Exerciţii: Arătaţi că pentru $n \to \infty$:

1) $\left(\dfrac{x^2}{1 + x^2}\right)^n \to 0$;

2) $\left(\dfrac{x}{1 + x^2}\right)^n \to 0$;

3) $\left(\dfrac{x^3}{4 + x^2}\right)^n$ tinde spre infinit pentru $x > 2$ şi spre 0 pentru $|x| < 2$.

3. Limita lui $\sqrt[n]{p}$

Şirul $a_n = \sqrt[n]{p}$, adică şirul $p, \sqrt{p}, \sqrt[3]{p}, \sqrt[4]{p}, \ldots$ are limita 1 pentru orice număr pozitiv fixat p:

(3) $$\sqrt[n]{p} \to 1 \text{ cînd } n \to \infty.$$

(Prin simbolul $\sqrt[n]{p}$ înțelegem, ca întotdeauna, rădăcina pozitivă de ordinul n. Pentru numere negative p nu există rădăcini reale de ordinul n, dacă n este număr par.)

Pentru a demonstra relația (3), trebuie să presupunem mai întîi că $p > 1$; atunci $\sqrt[n]{p}$ va fi și el mai mare decît 1. Astfel, putem pune

$$\sqrt[n]{p} = 1 + h_n,$$

unde h_n este o cantitate pozitivă, care depinde de n. Inegalitatea (2) arată atunci că

$$p = (1 + h_n)^n > nh_n.$$

Împărțind cu n vedem că

$$0 < h_n < \frac{p}{n}.$$

Deoarece șirurile $b_n = 0$ și $c_n = p/n$ au ambele limita 0, din raționamentul făcut în secțiunea 1 rezultă că și h_n are limita 0 atunci cînd n crește, și afirmația de mai sus este demonstrată pentru $p > 1$. Avem aici un exemplu tipic în care o relație de trecere la limită, în cazul nostru $h_n \to 0$, se obține incluzînd pe h_n între două margini, ale căror limite se obțin cu mai multă ușurință.

În treacăt fie zis, am obținut o evaluare pentru diferența h_n dintre $\sqrt[n]{p}$ și 1; această diferență trebuie să fie întotdeauna mai mică decît p/n.

Dacă $0 < p < 1$, atunci $\sqrt[n]{p} < 1$ și putem pune

$$\sqrt[n]{p} = \frac{1}{1 + h_n},$$

unde h_n este din nou un număr pozitiv, care depinde de n. Rezultă că

$$p = \frac{1}{(1 + h_n)^n} < \frac{1}{nh_n},$$

astfel încît

$$0 < h_n < \frac{1}{np}.$$

De aici rezultă că h_n tinde spre 0 cînd n crește. Deci, deoarece $\sqrt[n]{p} = 1/(1 + h_n)$, deducem că $\sqrt[n]{p} \to 1$.

Efectul egalizator al extragerii rădăcinii de ordinul n, care tinde să împingă orice număr pozitiv spre 1, atunci cînd n crește, este destul de puternic, astfel încît el să aibă același efect și în unele cazuri în care cantitatea de sub radical nu rămîne constantă. Vom demonstra că șirul $1, \sqrt{2}, \sqrt[3]{3}, \sqrt[4]{4}, \sqrt[5]{5}, \ldots$ tinde spre 1, adică

$$\sqrt[n]{n} \to 1$$

cînd n crește. Printr-un mic artificiu, acest lucru poate fi arătat cu ajutorul inegalității (2). În loc de a extrage rădăcina de ordinul n din n, vom extrage rădăcina de ordinul n din $\sqrt{n}$. Dacă punem $\sqrt[n]{\sqrt{n}} = 1 + k_n$, unde k_n este un număr pozitiv care depinde de n, atunci inegalitatea dă $\sqrt{n} = (1 + k_n)^n > nk_n$, astfel încît

$$k_n < \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}.$$

Deci

$$1 < \sqrt[n]{n} = (1 + k_n)^2 = 1 + 2k_n + k_n^2 < 1 + \frac{2}{\sqrt{n}} = \frac{1}{n}.$$

Membrul drept al acestei inegalități tinde spre 1 cînd n crește, astfel încît și $\sqrt[n]{n}$ trebuie să tindă spre 1.

4. Funcții discontinue ca limite de funcții continue

Putem considera limite de șiruri a_n, cînd a_n nu este un număr fixat, ci depinde de o variabilă x: $a_n = f_n(x)$. Dacă acest șir converge cînd $n \to \infty$, atunci și limita sa este o funcție de x:

$$f(x) = \lim f_n(x).$$

Astfel de reprezentări ale funcțiilor $f(x)$ ca limite de alte funcții sînt adesea folositoare pentru reducerea funcțiilor „superioare" $f(x)$ la funcții elementare $f_n(x)$.

Acest lucru este adevărat în particular pentru reprezentarea funcțiilor discontinue prin formule explicite. De exemplu, să considerăm șirul $f_n(x) = \frac{1}{1 + x^{2n}}$. Pentru $|x| = 1$ avem $x^{2n} = 1$ și deci $f_n(x) = 1/2$ pentru orice n,

astfel încît $f_n(x) \to 1/2$. Pentru $|x| < 1$ avem $x^{2n} \to 0$ şi deci $f_n(x) \to 1$, în timp ce pentru $|x| > 1$ avem $x^{2n} \to \infty$, şi deci $f_n(x) \to 0$. Rezumînd

$$f(x) = \lim \frac{1}{1 + x^{2n}} = \begin{cases} 1 \text{ pentru } |x| < 1, \\ \frac{1}{2} \text{ pentru } |x| = 1, \\ 0 \text{ pentru } |x| > 1. \end{cases}$$

Deci funcţia discontinuă $f(x)$ este reprezentată ca limită a unui şir de funcţii raţionale continue.

Un alt exemplu interesant de natură similară este dat de şirul

$$f_n(x) = x^2 + \frac{x^2}{1 + x^2} + \frac{x^2}{(1 + x^2)^2} + \ldots + \frac{x^2}{(1 + x^2)^n}.$$

Pentru $x = 0$, toate valorile $f_n(x)$ sînt nule şi deci $f(0) = \lim f_n(0) = 0$. Pentru $x \neq 0$, expresia $1/(1 + x^2) = q$ este pozitivă şi mai mică decît 1; rezultatele referitoare la seria geometrică garantează convergenţa lui $f_n(x)$ pentru $n \to \infty$. Limita, adică suma seriei geometrice infinite, este $\dfrac{x^2}{1 - q} = \dfrac{x^2}{1 - \dfrac{1}{1 + x^2}}$, care este egală cu $1 + x^2$. Astfel, vedem că $f_n(x)$ tinde spre funcţia $f(x) = 1 + x^2$ pentru $x \neq 0$ şi spre $f(x) = 0$ pentru $x = 0$. Această funcţie are o discontinuitate aparentă în $x = 0$.

*5. Limite prin iterare

Adesea termenii unui şir sînt astfel, încît a_{n+1} se obţine din a_n prin acelaşi procedeu prin care a_n s-a obţinut din a_{n-1}; acelaşi proces, repetat indefinit, produce întregul şir pornind de la termenul iniţial. În astfel de cazuri, vorbim despre un proces de „iterare”.

De exemplu, şirul

$$1, \sqrt{1 + 1}, \sqrt{1 + \sqrt{2}}, \sqrt{1 + \sqrt{1 + \sqrt{2}}}, \ldots$$

are o astfel de lege de formare; fiecare termen care urmează după primul se formează luînd rădăcina pătrată din 1 plus termenul precedent. Astfel formula

$$a_1 = 1, \; a_{n+1} = \sqrt{1 + a_n}$$

defineşte întregul şir. Să găsim limita lui. Evident, a_n este mai mare decît 1 pentru $n > 1$. Mai mult, a_n este un şir crescător pentru că

$$a_{n+1}^2 - a_n^2 = (1 + a_n) - (1 + a_{n-1}) = a_n - a_{n-1}.$$

Deci ori de cîte ori $a_n > a_{n-1}$ va rezulta că $a_{n+1} > a_n$. Știm însă că $a_2 - a_1 = \sqrt{2} - 1 > 0$, de unde rezultă prin inducție matematică că $a_{n+1} > a_n$ pentru orice n, astfel încît șirul este strict crescător. Mai mult, el este mărginit, pentru că, în baza rezultatelor precedente, avem

$$a_{n+1} = \frac{1 + a_n}{a_{n+1}} < \frac{1 + a_{n+1}}{a_{n+1}} = 1 + \frac{1}{a_{n+1}} < 2.$$

Din principiul șirurilor monotone deducem că pentru $n \to \infty$ avem $a_n \to a$, unde a este un număr cuprins între 1 și 2. Vedem cu ușurință că a este rădăcina pozitivă a ecuației pătratice $x^2 = 1 + x$. Într-adevăr, cînd $n \to \infty$, ecuația $a_{n+1}^2 = 1 + a_n$ devine $a^2 = 1 + a$. Rezolvînd această ecuație, găsim că rădăcina pozitivă este $a = \frac{1 + \sqrt{5}}{2}$. Astfel, putem rezolva această ecuație pătratică printr-un proces de iterare, care dă valoarea rădăcinii cu orice grad de aproximație dorit, dacă continuăm calculele destul de mult.

Putem rezolva multe alte ecuații algebrice prin iterare într-un mod asemănător. De exemplu, putem scrie ecuația cubică $x^3 - 3x + 1 = 0$ sub forma

$$x = \frac{1}{3 - x^2}$$

și apoi, alegînd o valoare oarecare pentru a_1, de pildă $a_1 = 0$ și definind

$$a_{n+1} = \frac{1}{3 - a_n^2},$$

obținem șirul $a_2 = 1/3 = 0{,}3333\ldots$, $a_3 = 9/26 = 0{,}3461\ldots$, $a_4 = 676/1947 = 0{,}3472\ldots$ etc. Se poate arăta că șirul a_n obținut în acest mod converge spre o limită $a = 0{,}3473\ldots$ care este o soluție a ecuației cubice date. Procesele de iterare ca acestea sînt deosebit de importante atît în matematica pură, în care ele furnizează „demonstrații de existență”, cît și în matematica aplicată, în care ele dau metode de aproximare pentru rezolvarea multor tipuri de probleme.

Exerciții asupra limitelor: Pentru $n \to \infty$:

1) Arătați că $\sqrt{n + 1} - \sqrt{n} \to 0$.

(Indicație: scrieți diferența sub forma

$$\frac{\sqrt{n + 1} - \sqrt{n}}{\sqrt{n + 1} + \sqrt{n}} (\sqrt{n + 1} + \sqrt{n}).)$$

2) Găsiți limita lui $\sqrt{n^2 + a} - \sqrt{n^2 + b}$.

3) Găsiți limita lui $\sqrt{n^2 + an + b} - n$.

4) Găsiți limita lui $\dfrac{1}{\sqrt{n+1} + \sqrt{n}}$.

5) Demonstrați că limita lui $\sqrt[n]{n+1}$ este 1.

6) Care este limita lui $\sqrt[n]{a^n + b^n}$ dacă $a > b > 0$?

7) Care este limita lui $\sqrt[n]{a^n + b^n + c^n}$ dacă $a > b > c > 0$?

8) Care este limita lui $\sqrt[n]{a^n b^n + a^n c^n + b^n c^n}$ dacă $a > b > c > 0$?

9) Vom vedea mai tîrziu (p.467) că $e = \lim (1 + 1/n)^n$. Care este atunci $\lim (1 + 1/n^2)^n$?

§ 2. EXEMPLU REFERITOR LA CONTINUITATE

Pentru a da o demonstrație precisă a continuității unei funcții este necesar să se verifice explicit definiția de la p. 328. Uneori acesta este un procedeu anevoios și de aceea este un fapt fericit că, așa cum vom vedea în cap. VIII, continuitatea este o consecință a derivabilității. Deoarece aceasta din urmă va fi stabilită sistematic pentru toate funcțiile elementare, putem merge pe drumul obișnuit al omiterii demonstrațiilor individuale dificile pentru demonstrarea continuității. Dar ca o altă ilustrare a definiției generale vom analiza un alt exemplu, și anume funcția $f(x) = \dfrac{1}{1+x^2}$. Putem restrînge pe x la un interval fixat $|x| \leqslant M$, unde M este un număr ales arbitrar. Scriind

$$f(x_1) - f(x) = \frac{1}{1+x_1^2} - \frac{1}{1+x^2} = \frac{x^2 - x_1^2}{(1+x^2)(1+x_1^2)} =$$
$$= (x - x_1)\frac{(x + x_1)}{(1+x^2)(1+x_1^2)}$$

găsim că pentru $|x| \leqslant M$ și $|x_1| \leqslant M$, avem:

$$|f(x_1) - f(x)| \leqslant |x - x_1||x + x_1| \leqslant |x - x_1| \cdot 2M.$$

Deci este clar că diferența din membrul stîng va fi mai mică decît orice număr pozitiv ε, cu condiția ca $|x_1 - x| < \delta = \dfrac{\varepsilon}{2M}$.

Ar trebui să remarcăm că sîntem foarte generoși în aprecierile noastre. Pentru valori mari ale lui x și x_1 cititorul va vedea cu ușurință că va fi suficient un număr δ cu mult mai mare.

Capitolul VII

MAXIME ŞI MINIME

INTRODUCERE

Un segment de dreaptă este drumul cel mai scurt care uneşte extremităţile lui. Un arc de cerc mare este cea mai scurtă curbă care uneşte două puncte de pe o sferă. Dintre toate curbele plane închise, de aceeaşi lungime, cercul cuprinde cea mai mare arie şi dintre toate suprafeţele închise de aceeaşi arie sfera cuprinde cel mai mare volum.

Proprietăţi de maxim şi minim de acest tip erau cunoscute grecilor, cu toate că rezultatele erau afirmate adesea fără o adevărată încercare de demonstraţie. Una dintre cele mai însemnate descoperiri ale grecilor este atribuită lui Heron, învăţatul din Alexandria, care a trăit în primul secol î.e.n. Se ştia de multă vreme că o rază de lumină pornită dintr-un punct P, care întîlneşte o oglindă plană L într-un punct R, este reflectată în direcţia unui punct Q, astfel încît PR şi QR formează unghiuri egale cu oglinda. Heron a găsit că dacă R' este un alt punct oarecare al oglinzii, atunci distanţa totală $PR' + R'Q$ este mai mare decît distanţa $PR + RQ$. Această teoremă, pe care o vom demonstra acum, caracterizează drumul real PRQ al luminii între P şi Q, ca fiind cel mai scurt drum între P şi Q, care atinge oglinda — descoperire care poate fi considerată ca germene al teoriei opticii geometrice.

Este foarte natural ca matematicienii să fie interesaţi în probleme de acest fel. În viaţa de toate zilele, probleme de maxim şi de minim, de „cel mai bun" şi „cel mai rău" apar mereu. Multe probleme de importanţă practică se prezintă sub această formă. De exemplu, cum trebuie construită o barcă, astfel încît ea să opună cea mai mică rezistenţă cînd se deplasează în apă? Ce dimensiuni trebuie să aibă o cutie cilindrică construită dintr-o anumită cantitate de material, pentru a avea un volum maxim?

Începînd din secolul al XVII-lea, teoria generală a extremelor, adică a maximelor şi minimelor, a devenit unul dintre principiile integratoare sistematice ale ştiinţei. Primii paşi făcuţi de Fermat în domeniul calculului diferenţial au fost determinaţi de dorinţa de a studia probleme de maxim şi

minim prin metode generale. În secolul care a urmat, aceste metode au fost substanțial îmbogățite o dată cu inventarea „calculului variațional". A devenit din ce în ce mai evident faptul că legile fizice ale naturii sînt exprimate în modul cel mai potrivit cu ajutorul unui principiu de minim, care permite o abordare naturală a soluției, mai mult sau mai puțin completă, a diferitelor probleme particulare. Una dintre cele mai remarcabile realizări ale matematicii contemporane este teoria valorilor staționare — o extindere a noțiunilor de maxim și minim, care se bazează concomitent pe analiză și topologie. Vom trata subiectul pe o cale foarte elementară.

§ 1. PROBLEME DE GEOMETRIE ELEMENTARĂ

1. Aria maximă a unui triunghi cu două laturi date

Sînt date două segmente a și b: se cere să se găsească triunghiul de arie maximă care are segmentele a și b ca laturi. Soluția este pur și simplu triunghiul dreptunghic, ale cărui catete sînt a și b. Într-adevăr, să considerăm orice triunghi care are pe a și pe b ca laturi, ca în fig. 176. Dacă h este înălțimea coborîtă pe baza a, atunci aria triunghiului este $A = \frac{1}{2}ah$. Însă $\frac{1}{2}ah$ este evident maximă atunci cînd h are valoarea cea mai mare, și aceasta se întîmplă atunci cînd h coincide cu b, adică pentru un triunghi dreptunghic. Deci aria maximă este $\frac{1}{2}ab$.

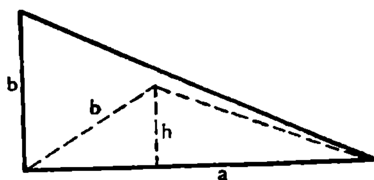


Fig. 176.

2. Teorema lui Heron. Proprietatea de extremum a razelor de lumină

Fiind date o dreaptă L și două puncte P și Q, de aceeași parte a lui L, cum trebuie ales punctul R pe dreapta L pentru ca suma $PR + RQ$ să dea cel mai scurt drum de la P la L și apoi la Q? Aceasta este problema lui Heron

a razei de lumină. (Dacă L ar fi malul unui rîu și dacă cineva ar trebui să meargă de la P la Q cît se poate de repede, aducînd o găleată cu apă de la rîu, atunci el ar avea de rezolvat chiar această problemă.) Pentru a găsi soluția, reflectînd pe P față de L ca într-o oglindă, obținem punctul P', astfel încît L să fie mediatoarea segmentului PP'. Dreapta $P'Q$ intersectează pe L în punctul cerut R. Este ușor de arătat că $PR + RQ$ este mai mic decît $PR' + R'Q$, pentru orice alt punct R' de pe L. Într-adevăr, $PR = P'R$ și $PR' = P'R'$; deci $PR + RQ = P'R + RQ = P'Q$ și $PR' + R'Q = P'R' + + R'Q$. Dar $P'R' + R'Q$ este mai mare decît $P'Q$ (pentru că suma a două laturi ale unui triunghi este mai mare decît a treia latură). Deci $PR' + R'Q$ este mai mare decît $PR + RQ$, ceea ce trebuia demonstrat. În cele ce urmează, presupunem că nici P și nici Q nu se află pe L.

Din fig. 177 vedem că $\sphericalangle 3 = \sphericalangle 2$ și $\sphericalangle 2 = \sphericalangle 1$, astfel încît $\sphericalangle 1 = \sphericalangle 3$. Cu alte cuvinte, R *este punctul pentru care* PR *și* RQ *fac unghiuri egale cu* L. De aici rezultă că o rază de lumină reflectată în L (dar prin reflexie, în baza experienței, unghiul de incidență este egal cu cel de reflexie) urmează într-adevăr drumul cel mai scurt de la P la L și apoi la Q, așa cum s-a afirmat mai sus.

Problema poate fi generalizată pentru a include mai multe drepte $L, M, \ldots$. De exemplu, să considerăm cazul în care avem două drepte L, M și două puncte P, Q situate ca în fig. 178, și problema de a găsi drumul de lungime

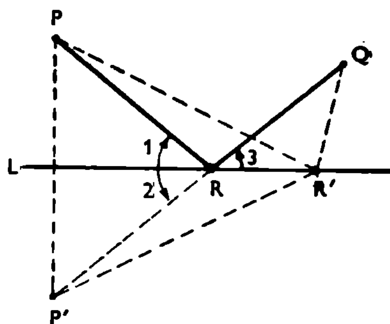


Fig. 177. Teorema lui Heron

minimă de la P la L, apoi la M și apoi la Q. Fie Q' simetricul lui Q față de M, iar Q'' simetricul lui Q' față de L. Să trasăm dreapta PQ'' care intersectează pe L în R, și dreapta RQ', care intersectează pe M în S; atunci R și S sînt punctele cerute, astfel încît $PR + RS + SQ$ este drumul de lungime minimă de la P la L, apoi de la L la M și apoi de la M la Q. Demonstrația acestui fapt este foarte asemănătoare cu aceea a problemei precedente și este

lăsată pe seama cititorului ca exercițiu. Dacă L și M ar fi oglinzi, o rază de lumină din P, reflectată de la L la M și apoi reflectată în Q ar întîlni pe L în R și pe M în S, deci raza de lumină ar parcurge drumul de lungime minimă.

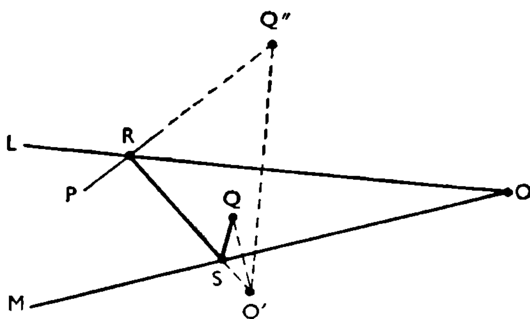


Fig. 178. Reflexia în două oglinzi

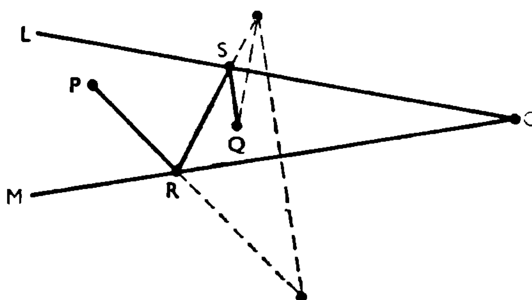


Fig. 179.

Am putea pune, de asemenea, problema găsirii drumului de lungime minimă care merge mai întîi de la P la M, apoi la L și de aici la Q. Aceasta ar da un drum $PRSQ$ (fig. 179), determinat într-un mod asemănător cu drumul precedent $PRSQ$. Lungimea primului drum poate fi mai mare, egală sau mai mică decît a celui de-al doilea.

* *Exercițiu:* Arătați că primul drum este mai scurt decît al doilea, dacă Q și R se află de aceeași parte a dreptei PQ. În ce caz cele două drumuri vor fi egale?

3. Aplicații la probleme asupra triunghiurilor

Cu ajutorul teoremei lui Heron se obțin cu ușurință soluțiile următoarelor două probleme.

a) Fiind date aria A și o latură $c = PQ$ ale unui triunghi, să se determine dintre toate aceste triunghiuri acela pentru care suma celorlalte două laturi

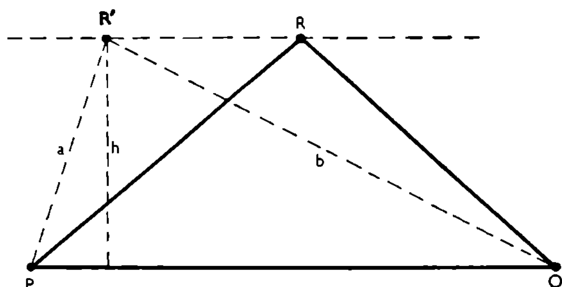


Fig. 180. Triunghiul de perimetru minim cu baza și aria date

a și b este minimă. Dînd de la început latura c și aria A a triunghiului înseamnă că a fost dată latura c și înălțimea h care cade pe c, deoarece $A = \frac{1}{2}hc$. Referindu-ne la fig. 180, problema se reduce deci la găsirea unui punct R, astfel încît distanța de la R la dreapta PQ să fie egală cu înălțimea h, și astfel încît suma $a + b$ să fie minimă. Din prima condiție rezultă că R trebuie să se găsească pe o dreaptă paralelă cu PQ și la o distanță h de aceasta. Răspunsul este dat cu ajutorul teoremei lui Heron, pentru cazul particular în care P și Q sînt la distanță egală de L: triunghiul PRQ cerut este isoscel.

b) Fiind date latura c și suma $a + b$ a celorlalte două laturi ale unui triunghi, să se determine dintre toate aceste triunghiuri acela care are aria cea mai mare. Aceasta este chiar reciproca problemei a). Soluția este din nou triunghiul isoscel pentru care $a = b$. După cum am arătat, acest triunghi realizează valoarea minimă pentru $a + b$, aria lui fiind dată. Cu alte cuvinte, oricare alt triunghi cu baza c și de aceeași arie are o valoare mai mare pentru $a + b$. Mai mult, din a) rezultă că orice triunghi cu baza c și cu o arie mai mare decît aceea a triunghiului isoscel are, de asemenea, o valoare mai mare pentru $a + b$. Deci orice alt triunghi cu aceleași valori pentru $a + b$ și c trebuie să aibă o arie mai mică, astfel încît triunghiul isoscel furnizează aria maximă, c și $a + b$ fiind date.

4. Proprietățile tangentelor la elipsă și hiperbolă. Proprietăți de extremum corespunzătoare

De teorema lui Heron sînt legate cîteva teoreme geometrice importante. Am demonstrat că dacă R este punctul de pe L, astfel încît $PR + RQ$ este minimă, atunci dreptele PR și RQ fac unghiuri egale cu L. Vom nota cu $2a$

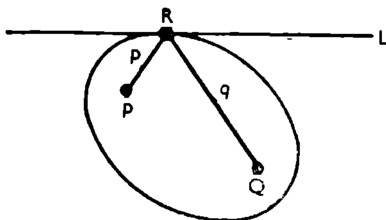


Fig. 181. Proprietatea tangentei la o elipsă

această distanță minimă. Fie p și q distanțele de la un punct oarecare din plan respectiv la punctele P și Q și să considerăm locul geometric al tuturor punctelor din plan pentru care $p + q = 2a$. Acest loc este o elipsă, care are focarele P și Q și care trece prin punctul R de pe dreapta L. Mai mult, ***L* trebuie să fie tangentă elipsei în** R. Dacă L ar mai intersecta elipsa într-un punct diferit de R, atunci ar exista un segment al lui L, conținut în interiorul elipsei; pentru orice punct al acestui segment, $p + q$ ar fi mai mic decît $2a$, deoarece se vede că $p + q$ este mai mic decît $2a$ în interiorul elipsei și mai mare decît $2a$ în exteriorul ei. Deoarece știm că $p + q \geqslant 2a$ pe L, acest lucru este imposibil. Deci L trebuie să fie tangentă la elipsă în R. Însă știm că PR și RQ fac unghiuri egale cu L, deci în treacăt am demonstrat următoarea teoremă importantă: o tangentă la o elipsă face unghiuri egale cu dreptele care unesc focarele cu punctul de tangență.

În strînsă legătură cu discuția precedentă se află următoarea problemă: fiind dată o dreaptă L și două puncte P și Q *de părți opuse* ale lui L (fig. 182), să se găsească un punct R pe L, astfel încît cantitatea $|p - q|$, adică valoarea absolută a *diferenței* distanțelor de la P și Q la R, să fie ***maximă***. (Vom presupune că L nu este mediatoarea segmentului BQ, deoarece atunci $p - q$ ar fi nulă pentru orice punct R de pe L și problema nu ar avea sens.) Pentru a rezolva această problemă, construind mai întîi simetricul lui P în raport cu L, obținem punctul P', de aceeași parte a lui L ca și Q. Pentru orice punct R' de pe L avem $p = R'P = R'P'$, $q = R'Q$. Deoarece R', Q și P' pot fi privite ca vîrfuri ale unui triunghi, cantitatea $|p - q| = |R'P' - R'Q|$ nu este niciodată mai mare decît $P'Q$, pentru că diferența dintre două laturi ale unui triunghi nu este niciodată mai mare decît a treia latură. Dacă R', P' și Q se află toate pe o aceeași dreaptă, atunci $|p - q|$ va fi *egal* cu $P'Q$,

după cum se vede din figură. De aceea, punctul R căutat este intersecția lui L cu dreapta care trece prin P' și Q. Ca și în cazul precedent, se vede cu ușurință că unghiurile pe care le fac RP și RQ cu L sînt egale, deoarece triunghiurile RPR' și $RP'R'$ sînt congruente.

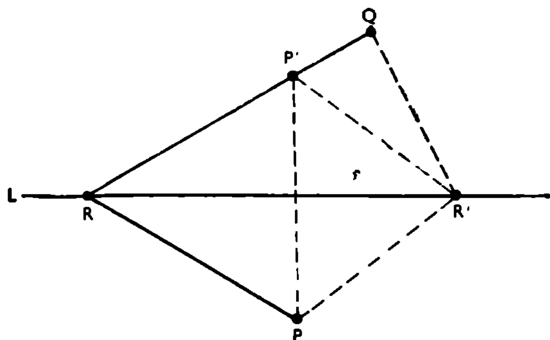


Fig. 182. $|PR - QR|$ = maxim

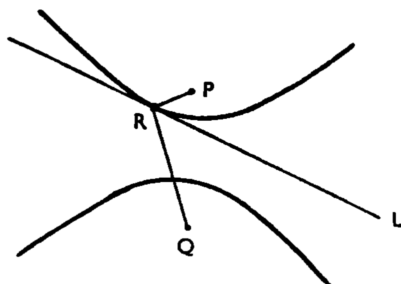


Fig. 183. Proprietatea tangentei la o hiperbolă

Această problemă este legată de o proprietate a tangentei la o hiperbolă, tot așa cum problema precedentă era în legătură cu elipsa. Dacă diferența maximă $|PR - QR|$ are valoarea $2a$, putem considera locul geometric al tuturor punctelor din plan pentru care $p - q$ are valoarea absolută $2a$. Acesta este o hiperbolă cu focarele în P și Q, care trece prin punctul R. După cum se vede cu ușurință, valoarea absolută a diferenței $p - q$ este mai mică decît $2a$, în regiunea cuprinsă între cele două ramuri ale hiperbolei, și mai mare decît $2a$, de acea parte a fiecărei ramuri în care se află focarul corespunzător. De aici rezultă, folosind în mod esențial același raționament ca și pentru elipsă, că L trebuie să fie tangentă la hiperbolă în R. Care anume

din cele două ramuri este atinsă de L, aceasta depinde de distanțele punctelor P și Q la dreapta L; dacă P este mai apropiat, ramura care înconjură pe P va atinge pe L; în mod asemănător pentru Q (fig. 183). Dacă P și Q sînt la distanță egală de L, atunci L nu va atinge nici o ramură a hiperbolei, dar va fi una dintre asimptotele curbei. Acest rezultat devine plauzibil, dacă observăm că în acest caz construcția precedentă nu va da un punct R la distanță finită, deoarece dreapta $P'Q$ va fi paralelă cu L.

În același mod ca și mai înainte, se demonstrează binecunoscuta teoremă: o tangentă la o hiperbolă într-un punct oarecare înjumătățește unghiul format de dreptele care unesc focarele cu punctul dat.

Ar putea să pară ciudat faptul că, dacă P și Q se află de aceeași parte a lui L, avem de rezolvat o problemă de minim, în timp ce dacă ele se află de o parte și de alta a lui L, trebuie să considerăm o problemă de maxim. Într-adevăr, în prima problemă fiecare din distanțele p, q, și deci și suma lor, crește nemărginit pe măsură ce ne îndepărtăm spre infinit pe L, în ambele sensuri. Deci ar fi imposibil să găsim o valoare maximă pentru $p + q$, și problema de *minim* este singura posibilitate. Lucrurile stau cu totul altfel în cel de-al doilea caz, în care P și Q se află de-o parte și de alta a lui L. Aici, pentru a evita confuzia, trebuie să facem o deosebire între diferența $p - q$, opusul ei $q - p$ și valoarea ei absolută $|p - q|$; cea din urmă trebuie făcută *maximă*. Înțelegem mai ușor situația dacă facem ca punctul R să se deplaseze în lungul dreptei L, ocupînd diferite poziții R_1, R_2, R_3, ... Există un punct pentru care diferența $p - q$ este nulă: intersecția mediatoarei segmentului PQ cu L. De aceea, acest punct dă un minim pentru valoarea absolută $|p - q|$. Dar de o parte a acestui punct, p este mai mare decît q, iar de cealaltă parte este mai mic; deci cantitatea $p - q$ este pozitivă de o parte a punctului și negativă de cealaltă parte. În consecință, $p - q$ nu este nici maximă nici minimă în punctul în care $|p - q| = 0$. Însă punctul care face maximă cantitatea $|p - q|$ este un extremum al lui $p - q$. Dacă $p > q$, avem un maxim pentru $p - q$; dacă $q > p$, avem un maxim pentru $q - p$ și deci un minim pentru $p - q$. Obținerea unui maxim sau a unui minim pentru expresia $p - q$ depinde de poziția celor două puncte date P și Q față de dreapta L.

Am văzut că nu există o soluție a problemei de maxim, dacă P și Q sînt la egală distanță de L, deoarece atunci dreapta $P'Q$ din fig. 182 este paralelă cu L. Acest lucru corespunde faptului că expresia $|p - q|$ tinde spre o limită, atunci cînd R tinde la infinit în lungul lui L, în oricare dintre cele două sensuri. Această valoare limită este egală cu lungimea proiecției perpendiculare s a lui PQ pe L (cititorul ar putea demonstra acest fapt ca exercițiu). Dacă P și Q sînt la distanță egală de L, atunci $|p - q|$ va fi întotdeauna mai mică decît această limită și nu există un maxim, deoarece pentru fiecare punct R putem găsi un altul mai îndepărtat, pentru care $|p - q|$ are o valoare mai mare, dar nu este niciodată egală cu s.

*5. Distanţe extreme la o curbă dată

Mai întîi vom determina *cea mai scurtă* şi *cea mai lungă distanţă* de la un punct P la o curbă dată C. Pentru simplificare vom presupune că C este o curbă simplă închisă cu tangentă în fiecare punct, ca în fig. 184. (Noţiunea de tangentă la o curbă este acceptată aici pe o bază intuitivă şi va fi analizată

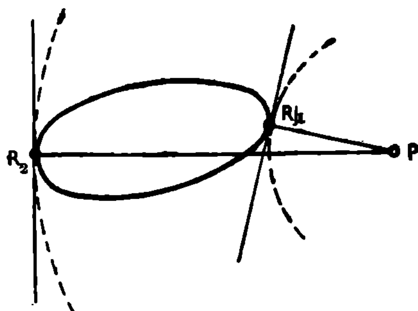


Fig. 184. Distanţe extreme la o curbă

în capitolul următor.) Răspunsul este foarte simplu: un punct R de pe C, pentru care distanţa PR are cea mai mică *sau* cea mai mare valoare, trebuie să fie astfel ales, încît dreapta PR să fie perpendiculară pe tangenta în R la C; cu alte cuvinte, PR este perpendiculară pe C. Demonstraţia se face în felul următor: cercul cu centrul în P, care trece prin R, trebuie să fie tangent curbei. Într-adevăr, dacă R este punctul de distanţă minimă, C trebuie să se afle în întregime în exteriorul cercului, şi de aceea nu îl poate traversa în R, în timp ce dacă R este punctul de distanţă maximă, C trebuie să se afle în întregime în interiorul cercului, şi din nou nu poate să îl traverseze în R. (Aceasta rezultă din faptul evident că distanţa de la orice punct la P este mai mică decît RP, dacă punctul se află în interiorul cercului şi este mai mare decît RP, dacă punctul se află în exteriorul cercului.) Deci cercul şi curba C vor fi tangente în punctul R. Dreapta PR, fiind rază a cercului, este perpendiculară pe tangenta la cerc în R, şi de aceea este perpendiculară pe C în R.

Facem observaţia că diametrul unei astfel de curbe închise C, adică cea mai lungă coardă a sa, trebuie să fie perpendiculară pe C la cele două extremităţi. Demonstraţia este lăsată ca exerciţiu pe seama cititorului. O afirmaţie asemănătoare ar trebui formulată şi demonstrată în trei dimensiuni.

Exerciţiu: Demonstraţi că cel mai scurt şi cel mai lung segment care uneşte două curbe închise care nu se intersectează este perpendicular pe cele două curbe la cele două extremităţi ale lui.

Problemele din secțiunea 4, referitoare la suma sau diferența distanțelor pot fi generalizate. Să considerăm în locul unei drepte L, o curbă simplă închisă C, cu tangentă în fiecare punct, și două puncte P și Q care nu se află pe C. Vrem să caracterizăm punctele de pe C, pentru care suma $p + q$ și diferența $p - q$ își ating valorile extreme, unde p și q reprezintă distanțele de la un punct oarecare al lui C, respectiv la P și Q. Nu putem folosi construcția simplă a punctelor simetrice, cu ajutorul căreia am rezolvat problemele în cazul în care C este o dreaptă, dar putem folosi proprietățile elipsei și ale hiperbolei pentru a rezolva aceste probleme. Deoarece C este o curbă închisă și nu mai este o curbă care se întinde la infinit, aici au sens atît problema de maxim, cît și cea de minim, pentru că putem fi siguri de faptul că cantitățile $p + q$ și $p - q$ au o cea mai mare *și* o cea mai mică valoare pe orice porțiune finită a unei curbe, și deci și pe o curbă închisă (cf. § 7).

În cazul sumei $p + q$, să presupunem că R este punctul de pe C, pentru care $p + q$ este maximă și fie $2a$ valoarea lui $p + q$ în R. Să considerăm elipsa cu focarele în P și Q, care este locul geometric al tuturor punctelor pentru care $p + q = 2a$. Această elipsă trebuie să fie tangentă la C în R (demonstrația este lăsată ca exercițiu pe seama cititorului). Am văzut însă că dreptele PR și QR fac unghiuri egale cu elipsa în R; deoarece elipsa este tangentă la C în R, dreptele PR și QR trebuie să facă și ele unghiuri egale cu C în R. Dacă $p + q$ este minimă pentru R, vedem în același mod că PR și QR fac unghiuri egale cu C în R. Astfel avem teorema: fiind dată o curbă închisă

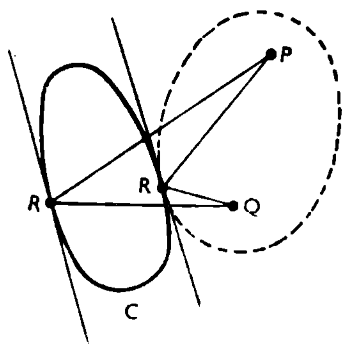


Fig. 185. Valorile maximă și minimă ale lui PR + QR

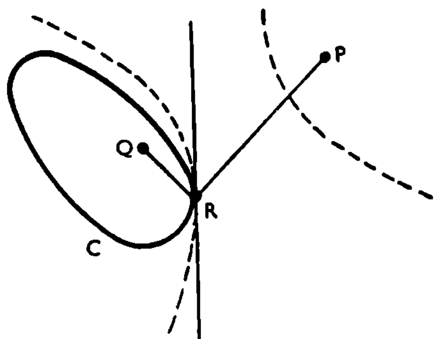


Fig. 186. Valoarea minimă a lui PR — QR

C și două puncte P și Q situate de aceeași parte a curbei C (ambele în exterior sau ambele în interiorul lui C) atunci într-un punct R al lui C, pentru care suma $p + q$ ia cea mai mare sau cea mai mică valoare de pe C, dreptele PR și QR fac unghiuri egale cu curba C (adică cu tangenta ei) în R.

Dacă P este în interiorul lui C, iar Q în exterior, această teoremă rămîne valabilă pentru cea mai mare valoare a lui $p + q$, dar încetează de a mai fi

adevărată pentru cea mai mică valoare, deoarece atunci elipsa degenerează într-o dreaptă.

Printr-un procedeu cu totul analog, folosind în locul proprietăților elipsei proprietățile hiperbolei, cititorul poate demonstra următoarea teoremă: fiind date o curbă închisă C și două puncte P și Q, situate de o parte și de alta a lui C, atunci într-un punct R al lui C, în care $p - q$ ia cea mai mare sau cea mai mică valoare pe C, dreptele PR și QR fac unghiuri egale cu C. Subliniem din nou faptul că problema pentru o curbă închisă C se deosebește de aceea referitoare la o curbă infinită, în măsura în care căutăm maximul valorii absolute $|p - q|$, deoarece pentru o curbă închisă există întotdeauna un maxim (ca și un minim) al lui $p - q$.

*§ 2. UN PRINCIPIU GENERAL CARE SE AFLĂ LA BAZA PROBLEMELOR DE EXTREMUM

1. Principiul

Problemele precedente sînt cazuri particulare ale unei probleme mai generale care se formulează cel mai bine în limbaj analitic. Dacă în problema găsirii valorilor extreme ale lui $p + q$ notăm cu x, y coordonatele punctului R, cu x_1, y_1 coordonatele punctului fixat P și cu x_2, y_2 pe acelea ale lui Q, atunci

$$p = \sqrt{(x - x_1)^2 + (y - y_1)^2}, \qquad q = \sqrt{(x - x_2)^2 + (y - y_2)^2},$$

și deci problema constă în găsirea valorilor extreme ale funcției

$$f(x, y) = p + q.$$

Această funcție este continuă peste tot în plan, însă punctul de coordonate x, y trebuie să se găsească pe curba C dată. Această curbă va fi definită printr-o ecuație $g(x,y) = 0$; de pildă, $x^2 + y^2 - 1 = 0$, dacă ea este cercul unitate. Problema constă atunci în găsirea valorilor extreme ale lui $f(x,y)$ cînd x și y sînt supuse la condiția ca $g(x,y) = 0$ și mai departe vom considera acest tip general de problemă.

Pentru a caracteriza soluțiile, vom considera familia de curbe $f(x,y) = c$, adică curbele date de ecuațiile de această formă, în care constanta c poate avea orice valoare, aceeași pentru toate punctele aceleiași curbe a familiei. Să presupunem că prin fiecare punct al planului sau cel puțin printr-o parte a lui, care conține curba C, trece o singură curbă din familia $f(x,y) = c$. Atunci pe măsură ce c variază, curba $f(x,y) = c$ va mătura o parte a planului și nici un punct din această parte nu va fi atins de două ori. (Curbele $x^2 - y^2 = c$, $x + y = c$ și $x = c$ sînt astfel de familii.) În particular, o curbă a familiei va trece prin punctul R_1 în care $f(x,y)$ ia cea mai mare valoare

de pe C, iar o altă curbă va trece prin punctul R_2, în care $f(x,y)$ ia cea mai mică valoare. Să notăm cu a cea mai mare valoare și cu b cea mai mică valoare. De o parte a curbei $f(x,y) = a$, valoarea lui $f(x,y)$ va fi mai mică decît a, iar de cealaltă parte va fi mai mare decît a. Deoarece $f(x,y) \leqslant a$ pe C, curba C trebuie să se afle în întregime de o parte a curbei $f(x,y) = a$, deci

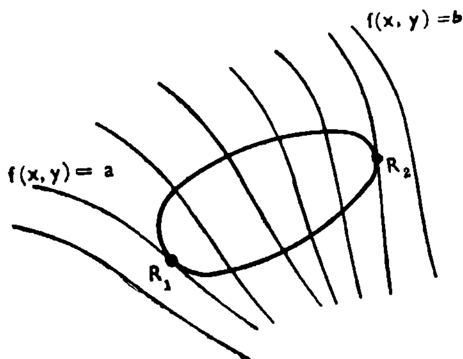


Fig. 187. Extremele unei funcții pe o curbă

ea trebuie să fie tangentă la această curbă în R_1. În mod asemănător, C trebuie să fie tangentă la curba $f(x,y) = b$ în R_2. Astfel am obținut teorema generală: *dacă într-un punct* R *al unei curbe* C *o funcție* f(x,y) *are o valoare extremă* a, *atunci curba* f(x,y) = a *este tangentă la* C *în* R.

2. Exemple

Rezultatele din paragraful precedent sînt, evident, cazuri particulare ale acestei teoreme generale. Dacă $p + q$ trebuie să aibă o valoare extremă, funcția $f(x,y)$ este egală cu $p + q$ și curbele $f(x,y) = c$ sînt elipse omofocale, cu focarele în P și Q. După cum rezultă din teorema generală, elipsele care trec prin punctele de pe C în care $f(x,y)$ ia valorile extreme, sînt tangente la C în aceste puncte. În cazul în care se caută extremele lui $p - q$, funcția $f(x,y)$ este egală cu $p - q$, curbele $f(x,y) = c$ sînt hiperbole omofocale cu focarele în P și Q, iar hiperbolele care trec prin punctele de extremum ale lui $f(x,y)$ sînt tangente la C.

Un alt exemplu este următorul: fiind dat un segment de dreaptă PQ și o dreaptă L care nu intersectează segmentul, care este punctul lui L din care segmentul PQ se vede sub cel mai mare unghi?

Funcția care trebuie maximizată în acest caz este unghiul θ, sub care se vede segmentul PQ din punctele lui L. Unghiul subîntins de PQ din orice

punct R din plan este o funcție $\theta = f(x,y)$ de coordonatele lui R. Din geometria elementară știm că familia de curbe $\theta = f(x,y) = c$ este familia de cercuri care trec prin P și Q, deoarece o coardă a unui cerc subîntinde același unghi în toate punctele circumferinței, aflate de aceeași parte a coardei. După cum se vede din fig. 190, numai două dintre aceste cercuri vor fi în

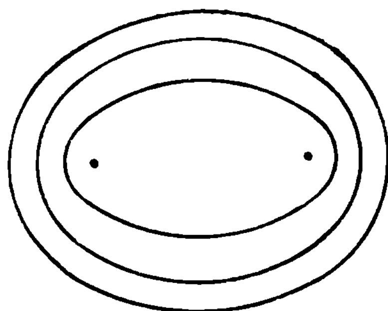


Fig. 188. Elipse omofocale

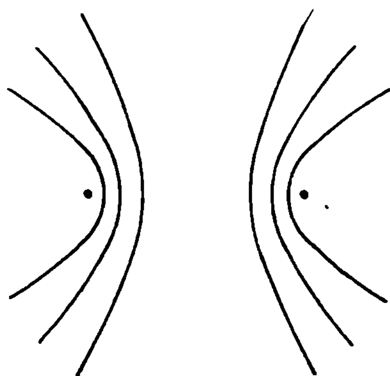


Fig. 189. Hiperbole omofocale

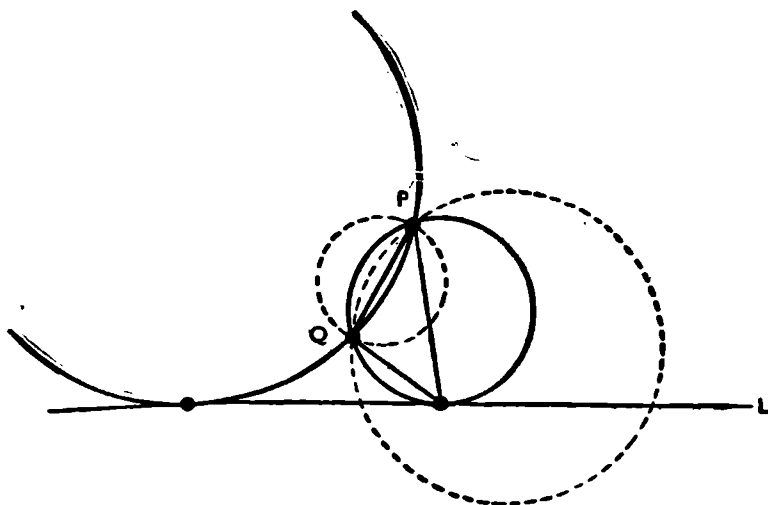


Fig. 190. Punctele de pe L, din care segmentul PQ *se vede sub un unghi maxim*

general tangente la L, centrele lor fiind de o parte și de alta a dreptei PQ. Unul dintre punctele de tangență dă un maxim absolut al lui θ, în timp ce celălalt punct dă un maxim „relativ" (adică valoarea lui θ va fi mai mică în orice punct aflat *într-o anumită vecinătate a punctului considerat*). Cel mai mare dintre cele două maxime, maximul absolut, este dat de acel punct de tangență care se află în unghiul ascuțit format de prelungirea lui PQ și de L, iar cel mai mic, de punctul care se află în unghiul obtuz format de aceste două drepte. (Punctul în care prelungirea segmentului PQ intersectează pe L dă valoarea minimă, egală cu zero, unghiului θ.)

Pentru a generaliza această problemă, putem înlocui pe L printr-o curbă C și putem căuta punctele R de pe C, din care un segment de dreaptă dat, PQ (care nu intersectează pe C) se vede sub cel mai mare sau cel mai mic unghi. Și în acest caz, cercul care trece prin P, Q și R trebuie să fie tangent curbei C în R.

§ 3. PUNCTELE STAȚIONARE ȘI CALCULUL DIFERENȚIAL

1. Extremele și punctele staționare

În raționamentele precedente nu a fost utilizată tehnica calculului diferențial. De fapt, metodele noastre elementare sînt mai simple și mai directe decît acelea ale analizei matematice. Ca regulă de gîndire științifică, este mai bine să se considere trăsăturile individuale ale unei probleme decît să se utilizeze în mod exclusiv metode generale, cu toate că eforturile individuale ar trebui să fie întotdeauna conduse de un principiu care clarifică înțelesul procedeelor speciale utilizate. Acesta este într-adevăr rolul calculului diferențial în problemele de extremum. Tendința modernă spre generalitate prezintă numai o latură a problemei, deoarece vitalitatea matematicii depinde în măsură decisivă de aspectul individual al problemelor și metodelor.

În dezvoltarea sa istorică, calculul diferențial a fost influențat puternic de probleme individuale de maxim și minim. Legătura dintre extreme și calculul diferențial rezultă din cele ce urmează. În cap. VIII vom face un studiu amănunțit al derivatei $f'(x)$ a unei funcții $f(x)$ și al înțelesului ei geometric. Pe scurt, derivata $f'(x)$ este panta tangentei la curba $y = f(x)$ în punctul (x,y). Este evident din punct de vedere geometric că într-un maxim sau un minim al unei curbe netede $y = f(x)$, tangenta la curbă trebuie să fie orizontală, adică panta ei trebuie să fie egală cu zero. Astfel, avem condiția $f'(x) = 0$ pentru valorile extreme ale lui $f(x)$.

Pentru a vedea ce înseamnă anularea lui $f'(x)$, să examinăm curba din fig. 191. Avem cinci puncte A, B, C, D, E în care tangenta la curbă este orizontală. Fie respectiv a, b, c, d, e valorile funcției $f(x)$ în aceste puncte. Maximul lui $f(x)$ în intervalul reprezentat este în D, minimul în A. Și punc-

tul B reprezintă un maxim, în sensul că pentru toate celelalte puncte dintr-o *anumită vecinătate* a lui B, $f(x)$ este mai mic decît b, cu toate că $f(x)$ este mai mare decît b, în puncte apropiate de D. Din acest motiv spunem că B este un *maxim relativ* al lui $f(x)$, în timp ce D este *maxim absolut*. În mod similar, C reprezintă un minim relativ, iar A minimul absolut. În sfîr-

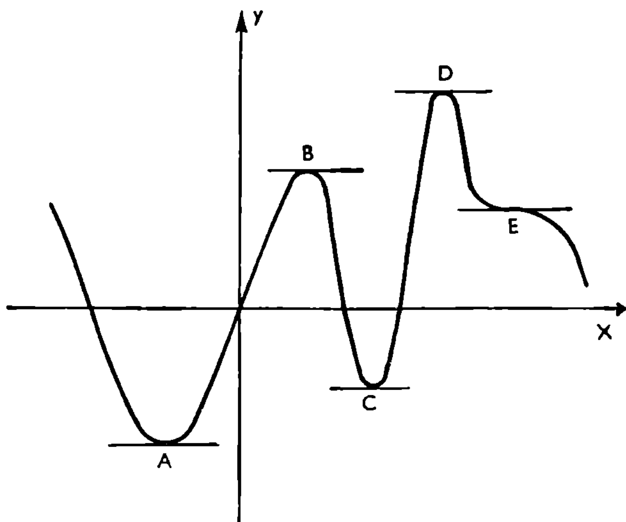


Fig. 191. Punctele staționare ale unei funcții

șit, în E nu există nici un maxim și nici un minim, cu toate că $f'(x) = 0$. De aici rezultă că anularea lui $f'(x)$ este o condiție *necesară*, dar nu *suficientă* pentru prezența unui extremum al unei funcții netede $f(x)$. Cu alte cuvinte, pentru orice extremum relativ sau absolut, $f'(x) = 0$, dar un punct în care $f'(x) = 0$ nu este în mod necesar un extremum. Un punct în care derivata se anulează, fie că este un extremum sau nu, se numește punct *staționar*. Printr-o analiză mai rafinată, se poate ajunge la condiții mai mult sau mai puțin complicate asupra derivatelor de ordin superior ale lui $f(x)$, care caracterizează pe deplin maximele, minimele și alte puncte staționare.

2. Maxime și minime ale funcțiilor de mai multe variabile. Puncte șa

Există probleme de maxim și minim, care nu pot fi exprimate cu ajutorul unei funcții $f(x)$ de o singură variabilă. Cel mai simplu caz de acest fel este acela al găsirii valorilor extreme ale unei funcții $z = f(x,y)$ de două variabile.

Putem reprezenta pe $f(x,y)$ prin cota z a unei suprafețe deasupra planului (x,y), pe care o putem interpreta, de pildă, ca peisaj muntos. Un maxim al lui $f(x,y)$ corespunde vîrfului unui munte, iar un minim, fundului unei depresiuni sau al unui lac. În ambele cazuri, dacă suprafața este netedă, planul tangent la suprafață va fi orizontal, însă mai există și alte puncte,

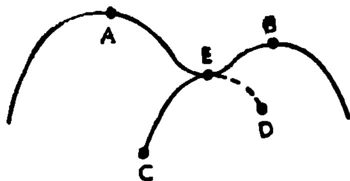


Fig. 192. O trecătoare

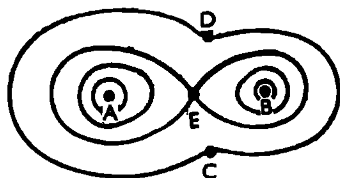


Fig. 193. Harta corespunzătoare cu curbe de nivel

pe lîngă vîrfuri de munți și funduri de văi, pentru care planul tangent este orizontal. Acestea sînt punctele date de trecători. Să examinăm aceste puncte în amănunt. Să considerăm ca în fig. 192 doi munți A și B alăturați și două puncte C și D de o parte și de alta a lor și să presupunem că vrem să trecem de la C spre D. Să considerăm mai întîi numai drumurile care duc de la C la D, care se obțin tăind suprafața cu un plan care trece prin C și D. Orice astfel de drum va avea un punct de înălțime maximă. Schimbînd poziția planului schimbăm drumul și va exista un drum CD pentru care înălțimea unui astfel de punct de înălțime maximă este *minimă*. Punctul E de înălțime maximă a acestui drum este o trecătoare, care în limbaj matematic se numește *punct șa*. Este evident că E nu este nici maxim, nici minim, deoarece putem găsi puncte oricît de apropiate de E, care sînt mai înalte sau mai joase decît E. În loc de a ne limita la drumuri care se află într-un plan, am putea considera tot atît de bine drumuri care nu sînt supuse acestei restricții și caracterizarea punctului șa E va rămîne neschimbată.

În mod asemănător, dacă dorim să trecem de la vîrful A la vîrful B, orice drum va avea un punct de altitudine minimă; dacă considerăm din nou numai secțiuni plane, va exista un drum AB, pentru care acest punct de altitudine minimă are altitudinea maximă și minimul acestui drum este tot punctul E găsit mai sus. Acest punct șa E are prin urmare proprietatea de a fi un minim de cea mai mare înălțime sau un maxim de cea mai mică înălțime, adică este un *maxi-minim* sau un *mini-maxim*. Planul tangent în E este orizontal, deoarece E fiind punctul minim al lui AB, dreapta tangentă la AB în E trebuie să fie orizontală, și în mod asemănător deoarece E este punctul maxim al lui CD, dreapta tangentă la CD în E trebuie să fie de asemenea orizontală. Planul tangent, care este planul determinat de aceste drepte, este de aceea orizontal. Astfel găsim trei tipuri diferite de puncte,

cu plane tangente orizontale: maxime, minime şi puncte şa; lor le corespund diferite tipuri de valori staționare ale lui $f(x,y)$.

Un alt mod de reprezentare a unei funcții $f(x,y)$ se realizează prin trasarea unor curbe de nivel ca acelea folosite în hărțile în care se indică altitudinile (cf. p. 304). O curbă de nivel este o curbă din planul (x,y), în lungul căreia funcția $f(x,y)$ are o valoare constantă; astfel, curbele de nivel coincid cu curbele familiei $f(x,y) = c$. Printr-un punct obişnuit al planului trece o singură curbă de nivel; un maxim sau un minim este înconjurat de curbe de nivel închise, în timp ce într-un punct şa se intersectează mai multe curbe de nivel. În fig. 193 sînt trasate curbele de nivel ale peisajului din fig. 192 şi proprietatea de maxi-minim a lui E este evidentă: orice drum care uneşte pe A cu B şi care nu trece prin E trebuie să treacă printr-o regiune în care $f(x,y) < f(E)$, în timp ce drumul AEB din fig. 192 are un minim în E. În acelaşi mod vedem că valoarea lui $f(x,y)$ în E este cel mai mic maxim pentru drumurile care unesc pe C cu D.

3. Punctele de minimax şi topologia

Există o legătură strînsă între teoria generală a punctelor staționare şi noțiunile topologiei. Aici putem da numai cîteva indicații scurte asupra acestor idei şi vom considera un singur exemplu.

Să considerăm peisajul muntos de pe o insulă B, de forma unui inel, cu frontierele C şi C'. Dacă reprezentăm din nou prin $u = f(x,y)$ altitudinea deasupra nivelului mării, cu $f(x,y) = 0$ pe C şi C' şi $f(x,y) > 0$ în interiorul

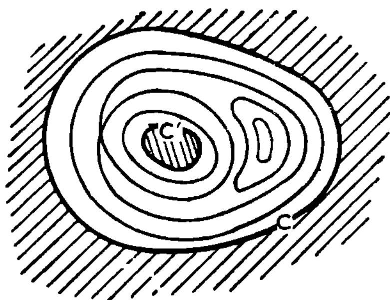


Fig. 194. Puncte staționare într-o regiune dublu conexă

lui B, atunci trebuie să existe cel puțin o trecătoare pe insulă, indicată în fig. 194 prin punctul în care liniile de nivel se intersectează. Intuitiv, acest lucru poate fi văzut dacă încercăm să mergem de la C la C', astfel încît să nu ne ridicăm la o altitudine mai mare decît este necesar. Orice drum de la C la C' trebuie să aibă un punct de altitudine maximă şi dacă alegem acel

drum al cărui punct de altitudine maximă este cît mai coborît cu putință, atunci punctul cel mai înalt al acestui drum este un punct șa al lui $u = f(x,y)$. (Există o excepție banală, cînd un plan orizontal este tangent crestei muntoase în lungul inelului.) Pentru un domeniu mărginit de p curbe trebuie să existe în general cel puțin $p - 1$ puncte staționare de tip minimax. Relații asemănătoare au fost descoperite de Marston Morse în spații cu mai multe dimensiuni, unde există o mai mare varietate de posibilități topologice şi de tipuri de puncte staționare. Aceste relații formează baza teoriei moderne a punctelor staționare.

4. Distanța de la un punct la o suprafață

Pentru distanța dintre un punct P şi o curbă închisă există (cel puțin) două valori staționare: un minim şi un maxim. Dacă încercăm să extindem acest rezultat la trei dimensiuni, nu apare nimic nou atîta timp cît considerăm o suprafață C topologic echivalentă cu o sferă, ca de pildă un elipsoid. Apar însă fenomene noi, dacă suprafaţa este de gen superior, ca de pildă un tor. Şi aici există o cea mai scurtă şi o cea mai lungă distanță de la P la un tor C, ambele segmente fiind perpendiculare pe C. În plus, găsim extreme de diferite tipuri, care reprezintă maxime de minime sau minime de maxime. Pentru a le găsi, să trasăm pe tor un cerc „meridian” închis L, ca în fig. 195, şi să căutăm pe L punctul Q cel mai apropiat de P. Apoi să încercăm să-l deplasăm pe L, astfel încît distanța PQ să devină : a) un minim. Punctul Q este pur şi simplu punctul de pe C, cel mai apropiat de P. b) un maxim. Acesta este un alt punct staționar. Tot atît de bine am putea

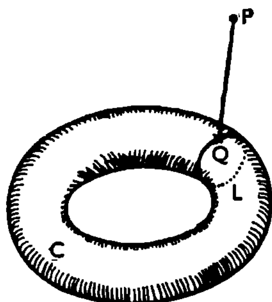


Fig. 195.

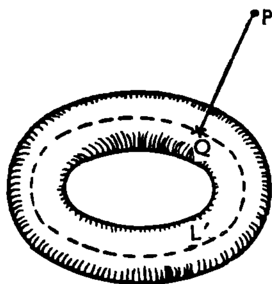


Fig. 196.

căuta pe L punctul cel mai îndepărtat de P şi apoi să găsim un cerc L, astfel încît această distanță maximă să fie: c) un maxim, care va fi atins în punctul de pe C cel mai îndepărtat de P. d) un minim. Obținem astfel patru valori staționare diferite ale distanței.

**Exercițiu:* Repetați raționamentul cu celălalt tip L' de curbă închisă de pe C care nu poate fi contractată la un punct, ca în fig. 196.

§ 4. PROBLEMA TRIUNGHIULUI LUI SCHWARZ

1. Demonstrația lui Schwarz

Hermann Amandus Schwarz (1843—1921) a fost un matematician distins al Universității din Berlin și unul dintre cei care au adus o mare contribuție la teoria modernă a funcțiilor și în analiză. El nu a disprețuit subiectele elementare și unul din ele tratează următoarea problemă: într-un triunghi ascuțitunghi dat, să se înscrie un alt triunghi, cu perimetrul minim (prin triunghi înscris înțelegem un triunghi care are fiecare vîrf pe cîte o latură a triunghiului inițial). Vom vedea că există un singur triunghi de acest fel și că vîrfurile lui sînt picioarele înălțimilor triunghiului dat. Acest triunghi se numește *triunghiul ortic.*

Schwarz a demonstrat proprietatea de minim a triunghiului ortic cu ajutorul metodei simetriei, folosind următoarea teoremă de geometrie elementară (fig.197): în fiecare vîrf P, Q, R al triunghiului ortic, cele două laturi ale triunghiului ortic fac unghiuri egale cu latura triunghiului inițial; acest unghi este egal cu unghiul din vîrful opus al triunghiului inițial. De exemplu, unghiurile ARQ și BRP sînt amîndouă egale cu C etc.

Pentru a demonstra această teoremă preliminară, să observăm că $OPBR$ este un patrulater inscriptibil, pentru că $\sphericalangle OPB$ și $\sphericalangle ORB$ sînt unghiuri drepte. În consecință $\sphericalangle PBO = \sphericalangle PRO$, deoarece ele subîntind același arc PO în cercul circumscris. Dar $\sphericalangle PBO$ este complementar $\sphericalangle C$, deoarece CBQ

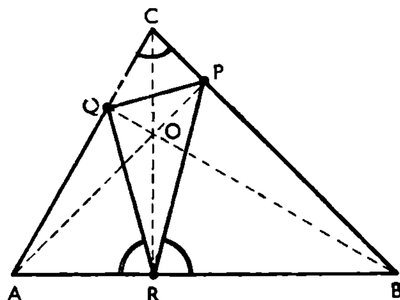


Fig. 197. Triunghiul ortic al lui ABC, *cu indicarea unghiurilor egale*

este un triunghi dreptunghic, iar $\sphericalangle PRO$ este complementar $\sphericalangle PRB$. De aceea, acest unghi este egal cu $\sphericalangle C$. În același mod, folosind patrulaterul $QORA$, vedem că $\sphericalangle QRA = \sphericalangle C$ etc.

Acest rezultat ne permite să enunțăm următoarea proprietate de simetrie a triunghiului ortic: deoarece $\sphericalangle AQR = \sphericalangle CQP$, simetrica dreptei RQ față

de latura *AC* este prelungirea lui *PQ* şi reciproc, obţinîndu-se un rezultat asemănător şi pentru celelalte laturi.

Vom demonstra acum proprietatea de minim a triunghiului ortic. Să considerăm, o dată cu triunghiul ortic, orice alt triunghi *UVW* înscris în triunghiul *ABC*. Să construim simetricul întregii figuri, mai întîi faţă de latura *AC* a triunghiului *ABC*, apoi să luăm simetricul triunghiului obţinut faţă de latura sa *AB*, apoi faţă de *BC*, din nou faţă de *AC* şi, în sfîrşit, faţă

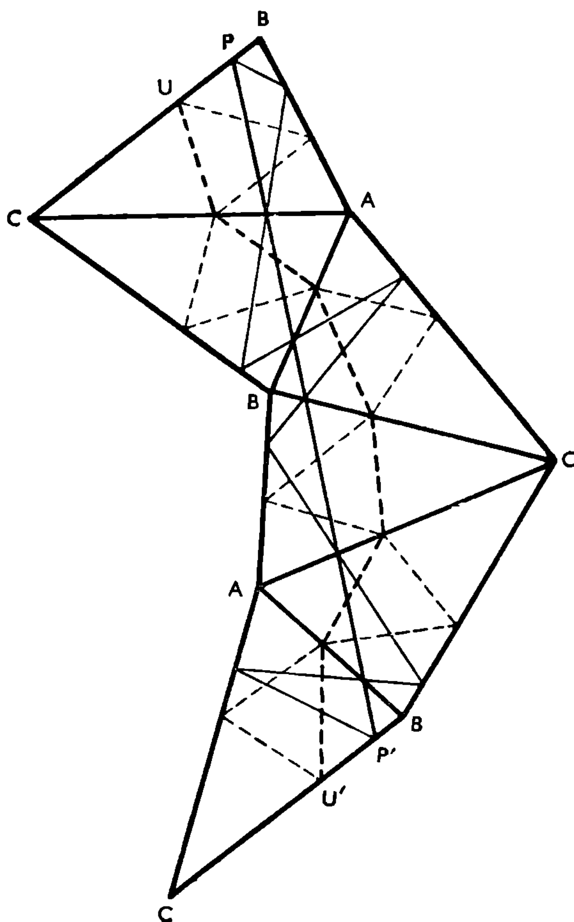


Fig. 198. Demonstraţia lui Schwarz a faptului că triunghiul ortic are perimetrul minim

de AB. În acest mod obținem în total șase triunghiuri egale, fiecare conținînd triunghiul ortic și celălalt triunghi înscris. Latura BC a ultimului triunghi este paralelă cu latura inițială BC. Într-adevăr, în prima simetrie latura BC este rotită în sensul acelor unui ceasornic, cu un unghi egal cu $2C$, apoi este rotită tot în sensul acelor unui ceasornic cu $2B$, iar în cea de-a treia simetrie ea nu este modificată; în cea de-a patra este rotită cu $2C$ în sensul opus acelor unui ceasornic, iar în cea de-a cincea cu $2B$, în sensul opus acelor unui ceasornic. Astfel, unghiul total de rotație este egal cu zero.

Datorită proprietății de simetrie a triunghiului ortic, segmentul de dreaptă PP' este egal cu dublul perimetrului triunghiului ortic; într-adevăr PP' se compune din șase porțiuni, egale pe rînd cu prima, a doua și a treia latură a triunghiului, fiecare apărînd de două ori. În mod asemănător, linia frîntă de la U la U' este egală cu dublul perimetrului celuilalt triunghi înscris. Această linie nu este mai scurtă decît segmentul rectiliniu UU'. Deoarece UU' este paralel cu PP', linia frîntă de la U la U' nu este mai scurtă decît PP', și de aceea perimetrul triunghiului ortic este cel mai mic dintre perimetrele tuturor triunghiurilor înscrise, ceea ce și trebuia demonstrat. Am stabilit, prin urmare, că există un minim și că el este dat de triunghiul ortic. Faptul că nu există un alt triunghi cu perimetrul egal cu acela al triunghiului ortic va rezulta din cele ce urmează.

2. O altă demonstrație

Poate cea mai simplă soluție a problemei lui Schwarz este următoarea, bazată pe faptul demonstrat mai înainte, că suma distanțelor de la două puncte P și Q la o dreaptă L este minimă în punctul R al lui L în care PR și QR fac unghiuri egale cu L, cu condiția ca P și Q să se afle de aceeași parte a lui L și nici unul dintre aceste puncte să nu se afle pe L. Să presupunem că triunghiul PQR, înscris în triunghiul ABC, rezolvă problema de minim. Atunci R trebuie să fie punctul de pe latura AB, pentru care $p + q$ este minimă și de aceea unghiurile ARQ și BRP trebuie să fie egale; în mod similar, $\sphericalangle AQR = \sphericalangle CQP$, $\sphericalangle BPR = \sphericalangle CPQ$. Astfel, triunghiul de perimetru minim, dacă există, trebuie să aibă proprietatea de egalitate a unghiurilor, folosită în demonstrația lui Schwarz. Rămîne de arătat că singurul triunghi cu această proprietate este triunghiul ortic. Mai mult, deoarece în teorema pe care se bazează această demonstrație se presupune că P și Q nu se află pe AB, demonstrația nu rămîne valabilă în cazul în care unul dintre punctele P, Q, R este un vîrf al triunghiului inițial (în care caz, triunghiul de perimetru minim ar degenera în dublul înălțimii corespunzătoare). Pentru a completa demonstrația, trebuie deci să arătăm că perimetrul triunghiului ortic este mai mic decît dublul oricărei înălțimi.

Pentru a trata prima chestiune, să observăm că dacă un triunghi înscris are proprietatea de egalitate a unghiurilor menționate mai sus, unghiurile

în P, Q şi R trebuie să fie respectiv egale cu $\sphericalangle A$, $\sphericalangle B$ şi $\sphericalangle C$. Într-adevăr, să presupunem, de pildă, că am avea $\sphericalangle ARQ = \sphericalangle C + \delta$. Atunci, deoarece suma unghiurilor unui triunghi este egală cu $180°$, unghiul din Q trebuie să fie egal cu $B - \delta$, iar cel din P cu $A - \delta$, pentru ca triunghiurile ARQ şi BRP să aibă suma unghiurilor egală cu $180°$. Însă atunci suma unghiurilor

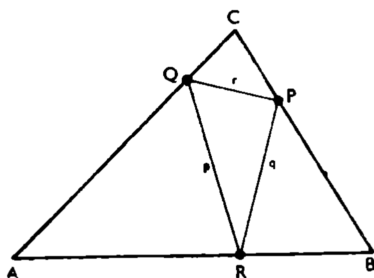


Fig. 199.

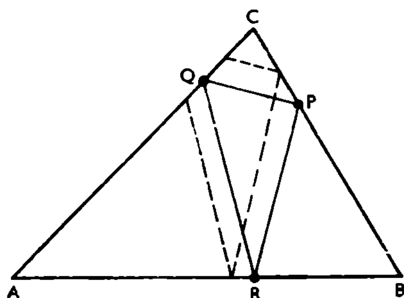


Fig. 200.

triunghiului CPQ este egală cu $A - \delta + B - \delta + C = 180° - 2\delta$; pe de altă parte, această sumă trebuie să fie egală cu $180°$, deci $\delta = 0$. Am văzut deja că triunghiul ortic are această proprietate de egalitate a unghiurilor. Orice alt triunghi cu această proprietate ar avea laturile paralele cu laturile corespunzătoare ale triunghiului ortic; cu alte cuvinte, el ar trebui să fie asemenea cu triunghiul ortic şi orientat în acelaşi mod. Cititorul poate arăta că în triunghiul dat nu poate fi înscris un alt triunghi cu aceste proprietăţi (fig. 200).

În sfîrşit, vom arăta că perimetrul triunghiului ortic este mai mic decît dublul oricărei înălţimi, cu condiţia ca unghiurile triunghiului iniţial să fie toate ascuţite. Ducem laturile QP şi QR şi coborîm perpendicularele din B pe QP, QR şi PR, obţinînd astfel punctele L, M şi N. Atunci QL şi QM sînt proiecţiile înălţimii QB, respectiv pe dreptele QP şi QR. Prin urmare, $QL + QM < 2QB$. Dar $QL + QM = p$, perimetrul triunghiului ortic. Într-adevăr, triunghiurile MRB şi NRB sînt egale, deoarece unghiurile MRB şi NRB sînt egale, iar unghiurile din M şi N sînt drepte. Deci $RM = RN$ şi $QM = QR + RN$. În acelaşi mod vedem că $PN = PL$, astfel încît $QL = QP + PN$. De aceea avem $QL + QM = QP + QR + PN + NR = QP + QR + PR = p$. Am arătat însă că $2QB > QL + QM$ şi deci p este mai mic decît dublul înălţimii QB; prin acelaşi raţionament, p este mai mic decît dublul oricărei înălţimi, după cum trebuia demonstrat. Deci, proprietatea de minim a triunghiului ortic este complet demonstrată.

Observăm, totodată, că construcţia precedentă permite un calcul direct al lui p. Ştim că unghiurile PQC şi RQA sînt egale cu B, şi de aceea $PQB = RQB = 90° - B$, astfel încît $\cos(PQB) = \sin B$. Folosind trigonometria elementară, obţinem $QM = QL = QB \sin B$ şi $p = 2QB \sin B$. În acelaşi mod se poate arăta că $p = 2PA \sin A = 2RC \sin C$. Din trigonometrie ştim că $RC\ a \sin B = b \sin A$ etc., ceea ce ne dă $p = 2a \sin B \sin C = 2b \sin C \sin A = 2c \sin A \sin B$. În sfîrşit, deoarece $a = 2r \sin A$, $b = 2r \sin B$, $c = 2r \sin C$, unde r este raza cercului circumscris, obţinem expresia simetrică $p = 4r \sin A \sin B \sin C$.

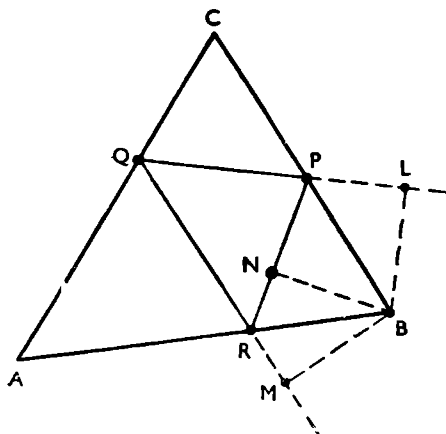


Fig. 201.

3. Triunghiuri obtuze

În cele două demonstraţii precedente s-a presupus că unghiurile A, B şi C sînt toate ascuţite. Dacă, de pildă, C este obtuz, ca în fig. 202, punctele P şi Q se vor afla în afara triunghiului. De aceea, triunghiul ortic nu mai poate fi considerat înscris în triunghi, în sensul strict al cuvîntului, afară

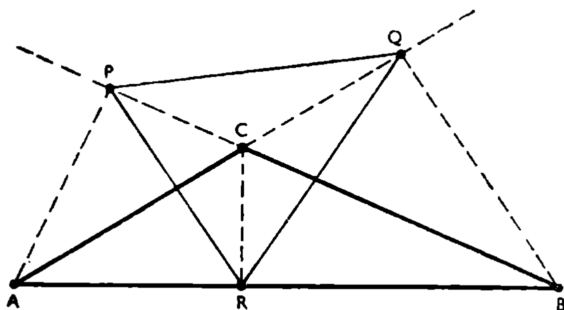


Fig. 202. Triunghiul ortic al unui triunghi obtuz

de cazul în care prin unghiul înscris înțelegem un triunghi ale cărui vîrfuri sînt pe laturile triunghiului inițial sau pe prelungirile lor. În orice caz, triunghiul ortic nu mai dă acum perimetrul minim, pentru că $PR > CR$ și $QR > CR$, deci $p = PR + QR + PQ > 2CR$. Deoarece raționamentul din prima parte a ultimei demonstrații arată că perimetrul minim, dacă nu este dat de triunghiul ortic, trebuie să fie egal cu dublul unei înălțimi, deducem că pentru unghiuri obtuze „triunghiul înscris" cu cel mai mic perimetru este cea mai scurtă înălțime, numărată de două ori, cu toate că aceasta nu este un triunghi propriu-zis. Totuși se observă că putem găsi un triunghi propriu, al cărui perimetru să difere de dublul înălțimii cu oricît de puțin vrem. Pentru cazul limită al triunghiului dreptunghic, cele două soluții — dublul înălțimii celei mai scurte și triunghiul ortic — coincid.

Problema interesantă, dacă triunghiul ortic are vreo proprietate de extremum pentru triunghiurile obtuze, nu poate fi discutată aici. Putem spune doar următoarele: triunghiul ortic nu dă un minim pentru suma laturilor $p + q + r$, ci o valoare staționară de tip minimax pentru expresia $p + q - r$, unde r este latura triunghiului înscris, opusă triunghiului obtuz.

4. Triunghiuri formate cu raze de lumină

Dacă triunghiul ABC reprezintă o cameră cu pereți reflectori, atunci triunghiul ortic este singurul drum triunghiular posibil parcurs de o rază de lumină. Nu sînt excluse alte drumuri închise parcurse de raze de lumină,

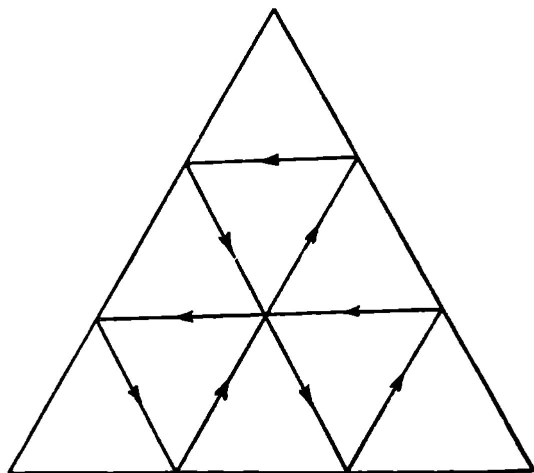


Fig. 203. Drum închis parcurs de o rază de lumină într-o oglindă triunghiulară

de forma unor poligoane, după cum arată fig. 203, însă triunghiul ortic este singurul poligon de acest fel, care are trei laturi.

Putem generaliza această problemă, căutînd „triunghiurile de lumină" posibile într-un domeniu arbitrar, mărginit de una sau mai multe curbe netede; adică căutăm triunghiurile care au vîrfurile undeva pe curbele frontieră și

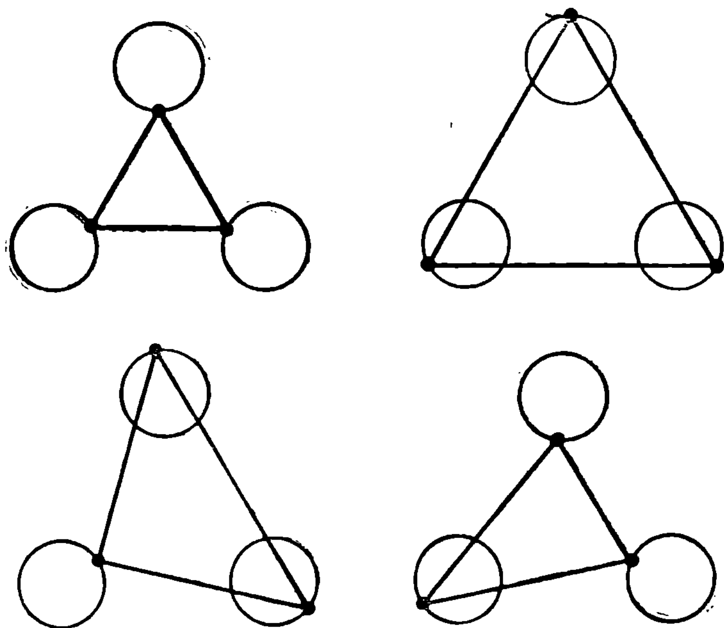


Fig. 204 — 207. Cele patru tipuri de triunghiuri de lumină care se formează între trei cercuri

sînt astfel, încît două laturi adiacente formează același unghi cu curba. După cum am văzut în § 1, egalitatea unghiurilor este o condiție atît pentru un maxim, cît și pentru un minim ale lungimii totale a celor două laturi, astfel încît, după împrejurări, putem găsi diferite tipuri de triunghiuri de lumină. De exemplu, dacă considerăm interiorul unei singure curbe netede C, atunci triunghiul înscris de perimetru maxim trebuie să fie un triunghi de lumină. Sau putem considera (la sugestia făcută autorilor de Marston Morse) exteriorul a trei curbe închise netede. Un triunghi de lumină ABC poate fi caracterizat prin faptul că perimetrul lui are o valoare staționară; această valoare poate fi un minim în raport cu toate cele trei puncte A, B, C, poate fi un minim în raport cu orice combinație de forma A și B și un maxim în raport cu al treilea punct C, poate fi un minim față de unul dintre

puncte şi un maxim faţă de celelalte două sau, în sfîrşit, poate fi un maxim faţă de toate cele trei puncte. În total este asigurată existenţa a cel puţin $2^3 = 8$ triunghiuri de lumină, deoarece pentru fiecare dintre cele trei puncte este posibil, în mod independent, fie un maxim, fie un minim.

*5. Observaţii referitoare la problemele de reflexie şi la mişcarea ergodică

O problemă de cel mai mare interes pentru dinamică şi optică este descrierea drumului sau a „traiectoriei" unei particule din spaţiu sau a unei raze de lumină, pentru un interval nelimitat de timp. Dacă printr-un artificiu fizic, particula sau raza este constrînsă să rămînă într-o porţiune mărginită a spaţiului, este deosebit de important să ştim dacă, la limită, traiectoria va umple spaţiul în întregime, cu o distribuţie aproximativ egală. O astfel de traiectorie se numeşte *ergodică*. Ipoteza existenţei ei este fundamentală pentru metodele statistice din dinamică şi teoria atomică modernă. Însă sînt cunoscute foarte puţine cazuri semnificative, în care se poate da o demonstraţie matematică riguroasă a „ipotezei ergodice".

Cele mai simple exemple se referă la cazul unei mişcări în interiorul unei curbe plane C, unde peretele C este presupus a avea proprietăţile unei oglinzi perfecte, care reflectă o particulă liberă sub acelaşi unghi sub care ea loveşte frontiera. De exemplu, o cutie dreptunghiulară (o masă de biliard idealizată, cu reflexie perfectă şi un punct material ca bila de biliard) duce în general la un drum ergodic; bila de biliard ideală, care îşi continuă drumul indefinit, va atinge vecinătatea oricărui punct, cu excepţia cîtorva poziţii iniţiale singulare şi a cîtorva direcţii. Omitem demonstraţia, cu toate că ea nu este dificilă în principiu.

De un deosebit interes este cazul unei mese eliptice, cu focarele F_1 şi F_2. Deoarece tangenta la o elipsă face unghiuri egale cu dreptele care unesc punctul de tangenţă cu cele două focare, orice traiectorie care trece printr-un focar va fi reflectată prin celălalt focar ş.a.m.d. Nu este greu de văzut că, indiferent de direcţia iniţială, după n reflexii, atunci cînd n tinde spre infinit, traiectoria tinde spre axa mare F_1F_2. Dacă raza iniţială nu trece printr-un focar, atunci există două posibilităţi. Dacă ea trece printre focare, atunci toate razele reflectate vor trece printre focare şi toate vor fi tangente unei anumite hiperbole care are pe F_1 şi pe F_2 ca focare. Dacă raza iniţială nu separă pe F_1 de F_2, atunci nici una dintre razele reflectate nu le vor separa şi ele vor fi toate tangente unei elipse care are pe F_1 şi pe F_2 ca focare. Astfel, în nici un caz mişcarea nu va fi ergodică pentru elipsă.

Exerciţii: 1) Demonstraţi că dacă raza iniţială trece printr-un focar al elipsei, atunci cea de-a n-a reflexie a razei iniţiale va tinde spre axa mare, cînd n tinde spre infinit.

2) Demonstraţi că dacă raza iniţială trece printre cele două focare, atunci toate razele reflectate vor face acelaşi lucru şi toate vor fi tangente unei hiperbole care are pe F_1 şi pe F_2 ca fo-

care; în mod asemănător, dacă raza inițială nu trece printre focare, nici una dintre razele reflectate nu va trece printre ele, și toate vor fi tangente unei elipse cu focarele în F_1 și F_2. (Indicație: Arătați că raza de dinaintea reflexiei și cea de după reflexia în R fac unghiuri egale, respectiv cu dreptele RF_1 și RF_2, și demonstrați apoi că tangentele la conice omofocale pot fi caracterizate în acest mod.)

§ 5. PROBLEMA LUI STEINER

1. Problema și soluția

O problemă foarte simplă, dar instructivă, a fost tratată de Jacob Steiner, celebru reprezentant al geometriei la Universitatea din Berlin, de la începutul secolului al XIX-lea. Trei localități A, B, C trebuie legate printr-un sistem de drumuri de lungime totală minimă. Din punct de vedere matematic, sînt date trei puncte A, B, C într-un plan și se caută un al patrulea punct P în plan, astfel încît suma $a + b + c$ să fie minimă, unde a, b, c sînt cele trei distanțe de la punctul P, respectiv la punctele A, B, C. Răspunsul la această problemă este următorul: dacă în triunghiul ABC, toate

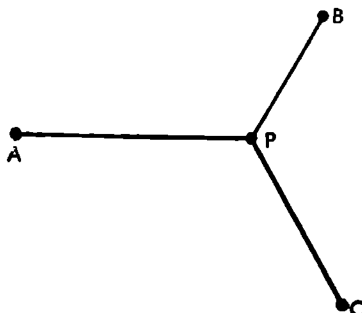


Fig. 208. Minimul sumei distanțelor la trei puncte

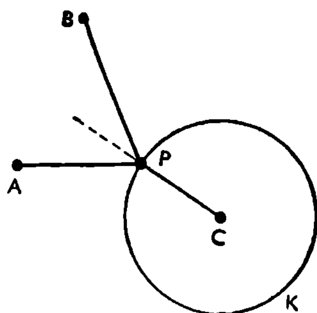


Fig. 209.

unghiurile sînt mai mici decît 120°, atunci P este punctul din care fiecare dintre cele trei laturi AB, BC, CA este văzută sub un unghi de 120°. Dacă însă un unghi al lui ABC, de exemplu unghiul din C, este egal sau mai mare decît 120°, atunci punctul P coincide cu vîrful C.

Este ușor de obținut această soluție, dacă folosim rezultatele precedente referitoare la extreme. Să presupunem că P este punctul de minim cerut. Există următoarele alternative: sau P coincide cu unul din vîrfurile A, B, C, sau P diferă de aceste vîrfuri. În primul caz, este clar că P trebuie să fie vîrful celui mai mare unghi C al triunghiului ABC, deoarece suma $CA + CB$

este mai mică decît oricare altă sumă formată cu două laturi ale triunghiului ABC. Astfel, pentru a completa demonstrația afirmației de mai sus, trebuie să analizăm al doilea caz. Fie K cercul de rază c, cu centrul în C. Atunci P trebuie să fie punctul de pe K, astfel încît $PA + PB$ să fie minimă. Dacă A și B sînt exterioare lui K, ca în fig. 209, atunci, conform rezultatului din § 1, PA și PB trebuie să formeze unghiuri egale cu cercul K și deci și cu raza PC, care este perpendiculară pe K. Deoarece același raționament se aplică, de asemenea, poziției lui P și cercului de rază a cu centrul în A, rezultă că toate cele trei unghiuri formate de PA, PB, PC sînt egale și deci sînt egale cu 120°, după cum s-a afirmat. Acest raționament a fost bazat pe ipoteza că A și B sînt ambele exterioare lui K, pe care trebuie s-o demonstrăm. Într-adevăr, dacă cel puțin unul dintre punctele A și B, de pildă A, ar fi pe K, sau în interiorul lui K, atunci ar trebui să avem $a + b \geqslant AB$, deoarece, în baza ipotezei P, nu coincide cu A sau cu B. Însă $AC \leqslant c$, deoarece A nu este în exteriorul lui K. Deci,

$$a + b + c \geqslant AB + AC,$$

ceea ce înseamnă că ar trebui să obținem cea mai mică sumă a distanțelor dacă P coincide cu A, contrar ipotezei noastre. Aceasta demonstrează că A și B sînt ambele exterioare cercului K. Proprietatea corespunzătoare se demonstrează în mod asemănător și pentru celelalte combinații: B, C față de un cerc de rază a cu centrul în A, și A, C față de un cerc de rază b, cu centrul în B.

2. Analiza alternativelor

Pentru a stabili care dintre cele două alternative pentru punctul P are loc în realitate, trebuie să examinăm construcția lui P. Pentru a-l găsi pe P, trebuie să trasăm doar cercurile K_1 și K_2 pe care două din laturi, de pildă AC și BC, subîntind arce de 120°. Atunci AC va subîntinde 120° din orice punct de pe arcul mai mic în care AC împarte pe K_1, însă va subîntinde 60° din orice punct de pe arcul mai lung. Intersecția celor două arce mai scurte, dacă există, va da punctul cerut P, pentru că nu numai corzile AC și BC vor subîntinde un unghi de 120° în punctul P, dar și AB va face același lucru, suma celor trei unghiuri fiind egală cu 360°.

Este limpede, din fig. 210, că dacă nici un unghi al triunghiului ABC nu este mai mare decît 120°, atunci cele două arce mai scurte se intersectează în interiorul triunghiului. Pe de altă parte, dacă un unghi C al triunghiului ABC este mai mare decît 120°, atunci cele două arce mai scurte ale lui K_1 și K_2 nu se vor intersecta, după cum se arată în fig. 211. În acest caz nu există nici un punct P din care toate cele trei laturi să fie văzute sub un unghi de 120°. Dar K_1 și K_2 determină prin intersecția lor un punct P', din care AC și BC subîntind fiecare unghiuri de 60°, în timp ce latura AB, opusă unghiului obtuz, subîntinde 120°.

Pentru un triunghi ABC, care are un unghi mai mare decît $120°$, nu există atunci nici un punct din care fiecare latură să fie văzută sub un unghi de $120°$. Deci punctul de minim P trebuie să coincidă cu un vîrf, deoarece

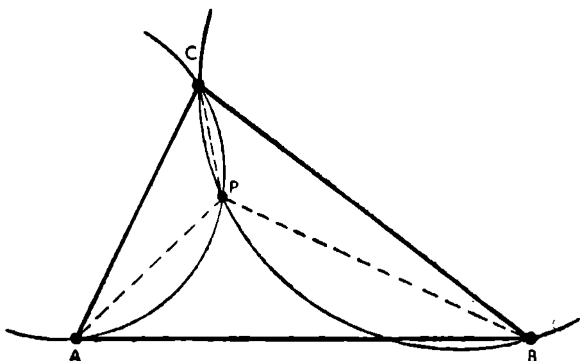


Fig. 210.

am arătat că aceasta este singura alternativă, și acest vîrf trebuie să fie vîrful unghiului obtuz. Dacă, pe de altă parte, toate unghiurile unui triunghi sînt mai mici decît $120°$, am văzut că se poate construi un punct P, din care fie-

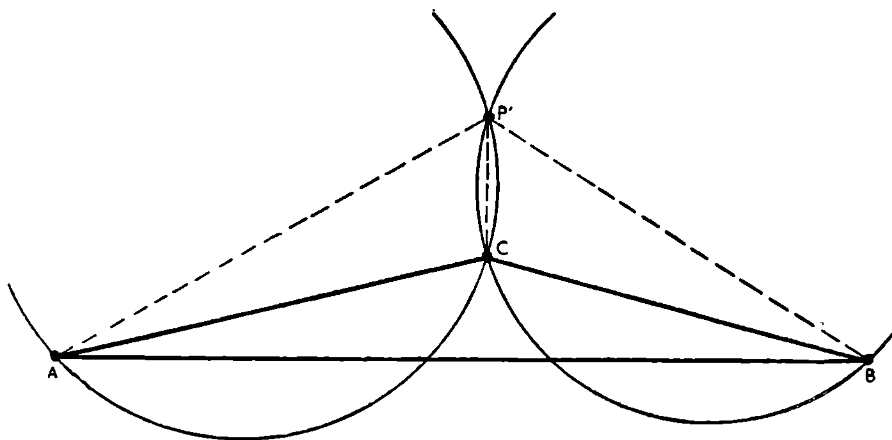


Fig. 211.

care latură să fie văzută sub un unghi de $120°$. Însă pentru a completa demonstrația acestei teoreme, trebuie să mai arătăm că $a + b + c$ va fi efectiv mai mică decît atunci cînd P coincide cu oricare din vîrfuri, deoarece am arătat

numai că P dă un minim, *dacă* lungimea totală minimă nu este atinsă într-unul din vîrfuri. În consecință, trebuie să arătăm că $a + b + c$ este mai mică decît suma oricăror două laturi, de pildă $AB + AC$. Pentru a face acest lucru, să prelungim pe BP și să proiectăm pe A pe această dreaptă,

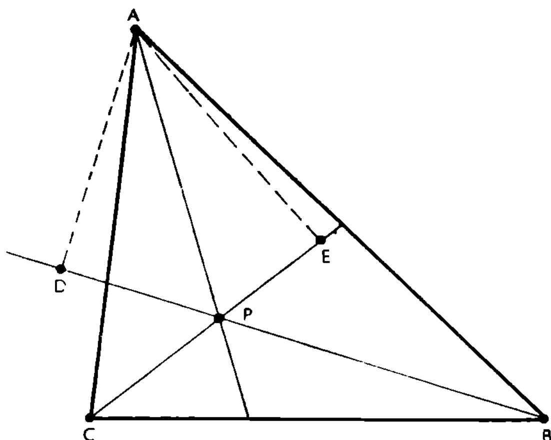


Fig. 212.

obținînd un punct D (fig. 212). Deoarece $\sphericalangle APD = 60°$, lungimea proiecției PD este egală cu $\frac{1}{2}a$. Dar BD este proiecția lui AB pe dreapta care trece prin B și prin P și, prin urmare, $BD < AB$. Însă $BD = b + \frac{1}{2}a$ și de aceea $b + \frac{1}{2}a < AB$. În același mod, proiectînd punctul A pe prelungirea lui PC, vedem că $c + \frac{1}{2}a < AC$. Prin adunare, obținem inegalitatea $a + b + c < AB + AC$. Deoarece știm că dacă punctul de minim nu este unul dintre vîrfuri, atunci el trebuie să coincidă cu P, rezultă, în sfîrșit, că P este efectiv punctul în care $a + b + c$ este minimă.

3. O problemă complementară

Metodele formale ale matematicii depășesc uneori intenția inițială. De exemplu, dacă unghiul C este mai mare decît 120°, construcția geometrică precedentă dă în locul soluției P (care, în acest caz, este punctul C),

un alt punct P', din care latura cea mai mare AB a triunghiului ABC apare sub un unghi de 120°, iar cele două laturi mai mici sub unghiuri de 60°. Desigur, P' nu rezolvă problema noastră de minim, dar putem bănui că are o legătură cu ea. Răspunsul este că P' rezolvă următoarea problemă:

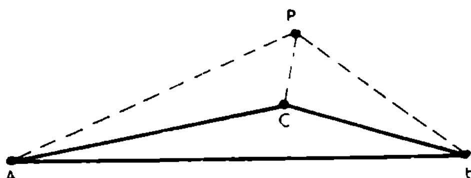


Fig. 213. $a + b - c = \text{minim}$

să se minimalizeze expresia $a + b - c$. Demonstrația este cu totul analogă aceleia date mai sus pentru $a + b + c$, bazată pe rezultatele din § 1, secțiunea 5, și este lăsată ca exercițiu pe seama cititorului. Combinînd acest rezultat cu rezultatul precedent avem teorema:

Dacă unghiurile unui triunghi ABC sînt mai mici decît 120°, atunci suma distanțelor a, b, c de la un punct oarecare, respectiv la A, B, C este minimă în punctul din care fiecare latură a triunghiului se vede sub un unghi de 120°, iar $a + b - c$ este minimă în vîrful C; dacă un unghi, de pildă C, este mai mare decît 120°, atunci $a + b + c$ este minimă în C, iar $a + b - c$ este minimă în punctul din care cele două laturi mai mici ale triunghiului sînt văzute sub unghiuri de 60°, iar latura cea mai mare este văzută sub un unghi de 120°.

Astfel, dintre cele două probleme de minim, una este rezolvată întotdeauna de construcția făcută cu ajutorul cercurilor, iar cealaltă de unul dintre vîrfuri. Pentru $\sphericalangle C = 120°$, cele două soluții ale fiecărei probleme, și chiar soluțiile celor două probleme coincid, deoarece punctul obținut prin construcție este chiar vîrful C.

4. Observații și exerciții

Dacă coborîm trei perpendiculare PA, PB, PC dintr-un punct P, interior unui triunghi echilateral UVW pe laturile lui, după cum se arată în fig. 214, atunci A, B, C și P vor forma figura studiată mai sus. Această observație poate folosi la rezolvarea problemei lui Steiner, pornind de la punctele A, B, C și găsind apoi punctele U, V, W.

Exerciții: 1) Efectuați demonstrația pe această cale, folosind faptul că din orice punct din interiorul unui triunghi echilateral, suma celor trei segmente perpendiculare pe laturi este constantă și egală cu înălțimea.

2) Folosind rezultatele corespunzătoare pentru cazul în care punctul P este exterior triunghiului UVW, discutați problema complementară.

În trei dimensiuni s-ar putea studia problema similară: fiind date patru puncte A, B, C, D, să se găsească un al cincilea punct P, astfel încît $a + b + c + d$ să fie minimă.

**Exercițiu:* Studiați această problemă și problema complementară prin metodele din § 1 sau folosind un tetraedru regulat.

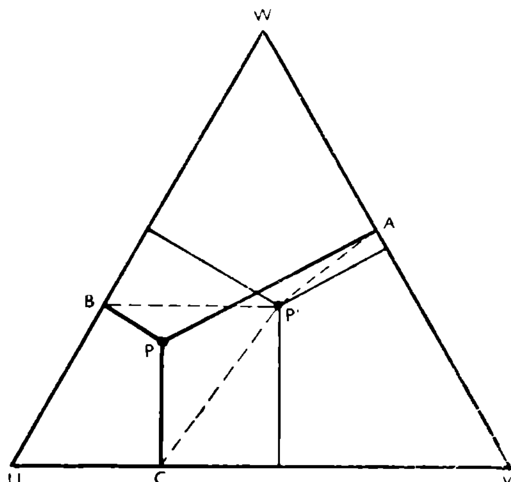


Fig. 214. O altă demonstrație a soluției lui Steiner

5. Generalizare la problema rețelei de drumuri

În problema lui Steiner sînt date trei puncte fixe A, B, C. Este natural să generalizăm această problemă, pentru cazul a n puncte date $A_1, A_2, \ldots, A_n$; căutăm punctul P din plan, pentru care suma distanțelor $a_1 + a_2 + \ldots + a_n$

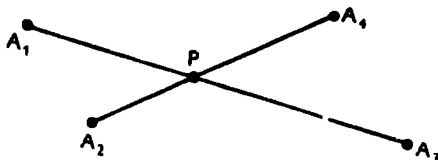


Fig. 215. Suma minimă a distanțelor la patru puncte

este minimă, unde a_i este distanța PA_i. (Pentru patru puncte dispuse ca în fig. 215, punctul P este punctul de intersecție al diagonalelor patrulaterului $A_1A_2A_3A_4$; cititorul poate demonstra acest lucru, ca exercițiu.) Această problemă, care a fost tratată de asemenea de Steiner, nu duce la rezultate intere-

sante. Ea constituie una dintre generalizările superficiale, întîlnite nu rareori în literatura matematică. Pentru a găsi extinderea într-adevăr semnificativă a problemei lui Steiner, trebuie să renunțăm la căutarea unui singur punct P, căutînd în schimb „rețeaua de drumuri" de lungime totală minimă. În enunț matematic aceasta înseamnă : *fiind date* n *puncte* $A_1, \ldots, A_n$, *să se găsească*

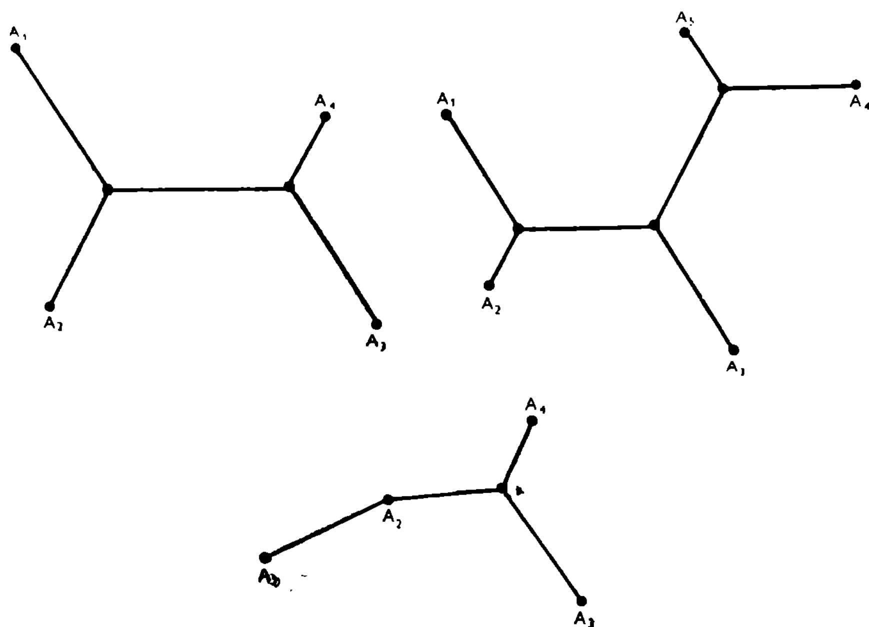


Fig. 216—218. Cele mai scurte rețele care unesc mai mult de trei puncte

un sistem conex de segmente de dreaptă, de lungime totală minimă, astfel încît oricare două dintre punctele date să poată fi unite printr-un poligon format din segmente ale rețelei.

Aspectul soluției va depinde, evident, de felul în care sînt dispuse punctele date. Cititorul poate studia cu folos acest subiect, cu ajutorul soluției problemei lui Steiner. Ne vom mulțumi aici să indicăm răspunsul pentru cazurile tipice arătate în figurile 216—218. În primul caz, soluția constă din cinci segmente cu două intersecții multiple, în care cîte trei segmente se întîlnesc formînd unghiuri de 120°. În al doilea caz, soluția conține trei intersecții multiple. Dacă punctele sînt dispuse în alt mod, este probabil ca astfel de figuri să nu mai poată fi construite. Una sau mai multe din intersecțiile multiple poate degenera și poate fi înlocuită prin unul sau mai multe din punctele date, ca în cel de-al treilea caz.

În cazul a n puncte date, vor exista cel mult $n - 2$ intersecții multiple, în fiecare dintre acestea întîlnindu-se trei segmente, care formează unghiuri de 120°.

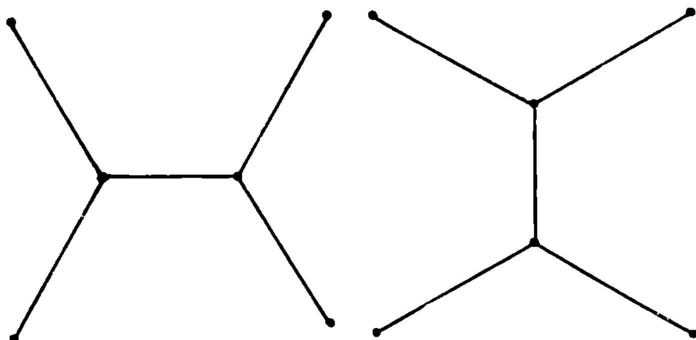


Fig. 219—220. Două rețele de lungime minimă care unesc patru puncte

Soluția problemei nu este determinată întotdeauna în mod unic. Pentru patru puncte A, B, C, D, care formează un pătrat, avem cele două soluții echivalente, indicate în fig. 219—220. Dacă punctele $A_1, A_2, \ldots, A_n$ sînt vîrfurile unui poligon simplu cu unghiuri destul de mari, atunci însuși poligonul va da soluția.

§ 6. EXTREME ȘI INEGALITĂȚI

Una dintre trăsăturile caracteristice ale matematicii superioare este rolul important jucat de inegalități. Rezolvarea unei probleme de maxim duce întotdeauna, în principiu, la o inegalitate, care exprimă faptul că cantitatea variabilă considerată este mai mică sau cel mult egală cu valoarea maximă furnizată de soluție. În multe cazuri, astfel de inegalități au un interes de sine stătător. Ca exemplu, vom considera inegalitatea importantă dintre media aritmetică și cea geometrică.

1. Media aritmetică și cea geometrică a două cantități pozitive

Începem cu o problemă simplă de maxim, care apare foarte des în matematica pură și în aplicațiile ei. În limbaj geometric, ea se reduce la următoarele: dintre toate dreptunghiurile cu perimetru dat, să se găsească dreptunghiul cu arie maximă. Soluția așa cum ne așteptăm, este dată de pătrat. Pentru a de-

monstra acest lucru, raționăm în modul următor. Fie $2a$ perimetrul dat al dreptunghiului. Atunci suma fixă a două laturi alăturate, de lungimi x și y, este $x+y$, în timp ce aria variabilă este xy, care trebuie făcută cît mai mare. „Media aritmetică" a lui x și y este

$$m = \frac{x+y}{2}.$$

Vom mai introduce și cantitatea

$$d = \frac{x-y}{2},$$

astfel încît

$$x = m + d, \quad y = m - d$$

și deci

$$xy = (m+d)(m-d) = m^2 - d^2 = \frac{(x+y)^2}{4} - d^2.$$

Deoarece d^2 este mai mare decît zero, cu excepția cazului în care $d = 0$, obținem imediat inegalitatea

$$(1) \qquad \sqrt{xy} \leqslant \frac{x+y}{2},$$

în care semnul egal apare numai dacă $d = 0$ și $x = y = m$.

Deoarece $x + y$ este fixată, rezultă că $\sqrt{xy}$, și deci aria xy, este maximă cînd $x = y$. Expresia

$$g = \sqrt{xy},$$

în care intervine rădăcina pătrată pozitivă, se numește „media geometrică" a cantităților pozitive x și y; inegalitatea (1) exprimă relația fundamentală dintre mediile aritmetică și geometrică.

Inegalitatea (1) rezultă de asemenea imediat din faptul că expresia

$$(\sqrt{x} - \sqrt{y})^2 = x + y - 2\sqrt{xy}$$

este în mod necesar nenegativă, fiind un pătrat, și este nulă numai dacă $x = y$.

Se poate da o demonstrație geometrică a inegalității, considerînd dreapta fixă $x + y = 2m$ din plan și familia de curbe $xy = c$, unde c este constantă pentru fiecare dintre aceste curbe (hiperbole) și variază de la curbă la curbă. Din

fig. 221 este evident că curba cu valoarea maximă a lui c, avînd un punct comun cu dreapta dată, va fi hiperbola tangentă la dreaptă în punctul $x = y = m$; pentru această hiperbolă avem deci $c = m^2$. Rezultă că

$$xy \leqslant \left(\frac{x+y}{2}\right)^2.$$

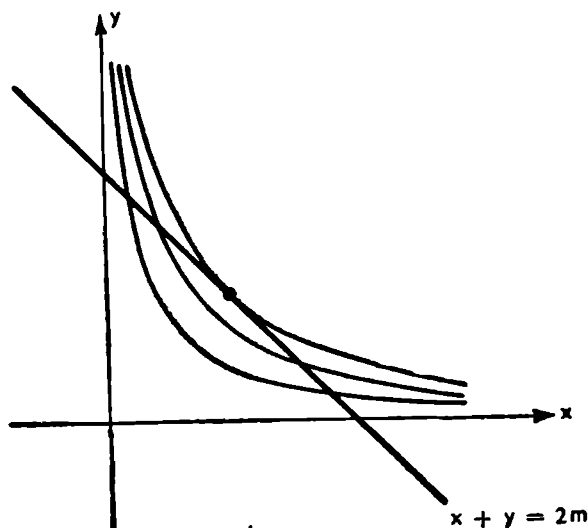


Fig. 221. Maximul lui xy *pentru* x + y *dat*

Ar trebui să remarcăm că orice inegalitate de forma $f(x, y) \leqslant g(x, y)$ poate fi citită în două moduri, și de aceea dă naștere unei proprietăți de maxim, ca și unei proprietăți de minim. De exemplu, inegalitatea (1) exprimă, de asemenea, faptul că dintre toate dreptunghiurile de arie dată, pătratul are perimetrul minim.

2. Generalizare la n variabile

Inegalitatea (1) dintre mediile aritmetică și geometrică a două cantități pozitive poate fi extinsă la orice număr de n cantități pozitive, notate cu $x_1, x_2, \ldots, x_n$. Numim numărul

$$m = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

media lor aritmetică, iar numărul

$$g = \sqrt[n]{x_1 x_2 \ldots x_n}$$

media lor geometrică, în care intervine rădăcina pozitivă de ordinul n. Teorema generală afirmă că

$$g \leqslant m, \tag{2}$$

și că $g = m$ numai dacă toate cantitățile x_i sînt egale.

Au fost imaginate numeroase demonstrații deosebite și ingenioase ale acestui rezultat general. Cea mai simplă cale este de a-l reduce la raționamentul folosit și în secțiunea 1, formulînd următoarea problemă de maxim: să se împartă o cantitate pozitivă dată C în n părți pozitive, $C = x_1 + \ldots + x_n$, astfel încît produsul $P = x_1 x_2 \ldots x_n$ să fie cît se poate de mare. Pornim de la ipoteza aparent evidentă, dar care va fi analizată mai tîrziu în § 7, că există un maxim al lui P și este atins de sistemul de valori

$$x_1 = a_1, \ \ldots, \ x_n = a_n.$$

Tot ceea ce trebuie să demonstrăm este că $a_1 = a_2 = \ldots = a_n$, pentru că în acest caz $g = m$. Să presupunem că acest lucru nu este adevărat, de pildă, că $a_1 \neq a_2$. Să considerăm cele n cantități

$$x_1 = s, \quad x_2 = s, \quad x_3 = a_3, \ldots, x_n = a_n,$$

unde

$$s = \frac{a_1 + a_2}{2}.$$

Cu alte cuvinte, înlocuim cantitățile a_i printr-un alt sistem, în care doar primele două sînt modificate și egalate, în timp ce suma totală C rămîne neschimbată. Putem scrie

$$a_1 = s + d, \quad a_2 = s - d,$$

unde

$$d = \frac{a_1 - a_2}{2}.$$

Noul produs este

$$P' = s^2 \cdot a_3 \ldots a_n,$$

în timp ce vechiul produs este

$$P = (s + d)(s - d) \cdot a_3 \ldots a_n = (s^2 - d^2) \cdot a_3 \ldots a_n,$$

astfel încît, în mod evident, dacă nu avem $d = 0$, atunci

$$P < P',$$

ceea ce contrazice ipoteza că P era valoarea maximă. Deci $d = 0$ şi $a_1 = a_2$. În acelaşi mod, putem demonstra că $a_1 = a_i$, unde a_i este oricare dintre numerele a, deci rezultă că toate numerele a sînt egale între ele. Deoarece $g = m$ atunci cînd toate numerele x_i sînt egale şi, deoarece am arătat că numai acestea dau valoarea maximă a lui g, rezultă că în caz contrar $g < m$, aşa cum se afirmă în teoremă.

3. Metoda celor mai mici pătrate

Media aritmetică a n numere $x_1, \ldots, x_n$, pe care nu mai trebuie să le presupunem pozitive în cele ce urmează, are o proprietate importantă de minim. Fie u o cantitate necunoscută, pe care vrem s-o determinăm cît se poate mai precis, cu ajutorul unui instrument de măsură. În acest scop facem n citiri, care pot furniza rezultate puţin diferite $x_1, \ldots, x_n$, datorită diferitelor surse de eroare experimentală. Atunci se pune problema care dintre valorile lui u trebuie să fie acceptată cu mai multă încredere? De obicei, ca valoare „adevărată" sau „optimă" se alege medie aritmetică $m = \dfrac{x_1 + x_2 + \ldots + x_n}{n}$. Pentru a da o justificare reală acestei ipoteze, trebuie să facem o discuţie detaliată de teoria probabilităţilor. Însă putem indica cel puţin o proprietate de minim a lui m, care îi justifică alegerea. Fie u o valoare posibilă a cantităţii măsurate. Atunci diferenţele $u - x_1, \ldots, u - x_n$ sînt abaterile acestei valori în diferitele citiri. Aceste abateri pot fi o parte pozitive, iar o parte negative şi este natural să presupunem că valoarea optimă a lui u este aceea pentru care abaterea totală este, într-un anumit sens, cît se poate de mică. Urmînd calea indicată de Gauss, de obicei sînt luate nu abaterile, ci pătratele lor $(u - x_i)^2$ ca măsuri potrivite ale impreciziei şi se alege ca valoare optimă dintre toate valorile posibile ale lui u, aceea care face ca suma pătratelor abaterilor

$$(u - x_1)^2 + (u - x_2)^2 + \ldots + (u - x_n)^2$$

să fie cît mai mică. *Această valoare optimă a lui u este tocmai media aritmetică* m, şi tocmai acest fapt constituie punctul de plecare în importanta „metodă a celor mai mici pătrate" a lui Gauss. Putem demonstra afirmaţia subliniată de mai sus printr-o metodă elegantă. Scriind

$$u - x_i = (m - x_i) + (u - m),$$

obţinem

$$(u - x_i)^2 = (m - x_i)^2 + (u - m)^2 + 2(m - x_i)(u - m).$$

Să adunăm acum toate aceste egalități pentru $i = 1, 2, \ldots, n$. Ultimii termeni dau $2(u - m)(nm - x_1 - \ldots - x_n)$, cantitate care este nulă în baza definiției lui m; prin urmare, obținem

$$(u - x_1)^2 + \ldots + (u - x_n)^2 = (m - x_1)^2 + \ldots + (m - x_n)^2 + n\,(m - u)^2.$$

Aceasta arată că

$$(u - x_1)^2 + \ldots + (u - x_n)^2 \geqslant (m - x_1)^2 + \ldots + (m - x_n)^2,$$

astfel încît semnul de egalitate are loc numai pentru $u = m$, care este tocmai ceea ce trebuia demonstrat.

Metoda generală a celor mai mici pătrate ia acest rezultat ca principiu conducător în cazuri mai complicate, în care problema este de a decide asupra unui rezultat plauzibil, pe baza unor măsurători ușor incompatibile. De exemplu, să presupunem că am măsurat coordonatele a n puncte, x_i, y_i care, teoretic vorbind, trebuie să se afle pe o dreaptă, și să presupunem că punctele obținute în acest mod empiric nu se află chiar pe o dreaptă. Cum trebuie să ducem dreapta, astfel încît ea să concorde cît mai bine cu cele n puncte observate? Rezultatul precedent sugerează următorul procedeu, care desigur ar putea fi înlocuit cu variante tot atît de rezonabile. Fie $y = ax + b$ ecuația dreptei, astfel încît problema constă în găsirea coeficienților a și b. Distanța în direcția y, de la punctul x_i, y_i la dreaptă, este dată de $y_i - (ax_i + b) = y_i - ax_i - b$, cu semn pozitiv sau negativ, după cum punctul este deasupra sau dedesubtul dreptei. Deci pătratul acestei distanțe este $(y_i - ax_i - b)^2$ și metoda constă în determinarea lui a și b, astfel încît expresia

$$(y_1 - ax_1 - b)^2 + \ldots + (y_n - ax_n - b)^2$$

să aibă cea mai mică valoare. Avem aici o problemă de minim, în care intervin două cantități necunoscute a și b. Discuția detaliată a soluției, deși este foarte simplă, va fi omisă.

§ 7. EXISTENȚA UNUI EXTREMUM. PRINCIPIUL LUI DIRICHLET

1. Observații generale

În cîteva din problemele precedente de extremum, s-a demonstrat direct că soluția dă cel mai bun rezultat din totalitatea posibilităților. Un exemplu frapant este soluția lui Schwarz a problemei triunghiului, în care am putut vedea că nici un triunghi înscris nu are un perimetru mai mic decît acela al triunghiului ortic. Alte exemple sînt problemele de minim sau de maxim, ale căror soluții depind de o inegalitate explicită, ca de pildă aceea dintre mediile aritmetică și geometrică. Însă în cîteva din problemele noastre am urmat o cale diferită. Am început cu ipoteza că a fost găsită o soluție, apoi am

analizat această ipoteză şi am tras concluzii care eventual permiteau o descriere şi o construcţie a soluţiei. Aşa s-a întîmplat, de pildă, cu soluţia problemei lui Steiner şi cu a doua tratare a problemei lui Schwarz. Cele două metode sînt logic diferite. Prima, într-un anumit fel este mai perfectă, deoarece dă o demonstraţie mai mult sau mai puţin constructivă a soluţiei. A doua metodă, aşa cum am văzut în cazul problemei triunghiului, poate fi mai simplă, însă nu este atît de directă, şi mai ales are o structură condiţională, deoarece porneşte de la ipoteza că *există* o soluţie a problemei. Ea dă soluţia numai dacă existenţa ei este asigurată sau demonstrată. Fără această ipoteză ea arată doar că *dacă* există soluţia, atunci ea trebuie să aibă o anumită structură[1].

Datorită evidenţei aparente a premisei că soluţia există, pînă tîrziu, în secolul al XIX-lea, matematicienii nu au dat nici o atenţie aspectului logic al problemei şi au admis existenţa unei soluţii pentru problemele de extremum, ca fiind un lucru de la sine înţeles. Unii dintre cei mai mari matematicieni din secolul al XIX-lea — Gauss, Dirichlet şi Riemann — au folosit această ipoteză fără a face o analiză mai atentă, ca bază pentru obţinerea unor teoreme profunde şi de altfel greu accesibile pe o altă cale, din fizica matematică şi teoria funcţiilor. Apogeul a avut loc în 1849, cînd Riemann şi-a publicat teza sa de doctorat asupra fundamentelor teoriei funcţiilor de variabilă complexă. Această lucrare concisă, una din marile realizări de pionierat ale matematicii moderne, era atît de neortodoxă în abordarea subiectului, încît multă lume ar fi vrut s-o ignore. Weierstrass era pe atunci cel mai proeminent matematician al Universităţii din Berlin şi conducătorul recunoscut în construirea unei teorii riguroase a funcţiilor. Impresionat, dar oarecum sceptic, el a descoperit curînd o lacună logică în lucrare, pe care autorul nu s-a îngrijit s-o elimiue. Critica zguduitoare a lui Weierstrass, deşi nu l-a tulburat pe Riemann, a avut ca prim rezultat o neglijare aproape generală a teoriei sale. Cariera meteorică a lui Riemann a avut un sfîrşit brusc cîţiva ani după aceea, cînd a murit de tuberculoză. Cu toate acestea, ideile sale au avut întotdeauna discipoli entuziaşti. După 50 de ani de la publicarea tezei sale, Hilbert a reuşit să deschidă calea unui răspuns complet la problemele pe care el le-a lăsat nerezolvate. Întreaga dezvoltare a matematicii şi a fizicii matematice a devenit unul dintre marile triumfuri din istoria ştiinţei contemporane.

În articolul lui Riemann, punctul vulnerabil este problema existenţei unui minim. Riemann şi-a bazat o mare parte a teoriei sale pe ceea ce el a numit principiul lui Dirichlet (Dirichlet a fost profesorul lui Riemann la Göttingen şi a predat acest principiu, pe care însă nu l-a scris niciodată). Să presupunem, de pildă, că o parte a unui plan sau a unei suprafeţe este acoperită cu poleială şi că o conectăm la o baterie electrică, producînd în acest mod un curent

[1] Necesitatea logică a stabilirii existenţei unui extremum este ilustrată de următoarea eroare: 1 este cel mai mare întreg. Într-adevăr, să notăm cu x cel mai mare întreg. Dacă $x > 1$, atunci $x^2 > x$, deci x nu ar fi cel mai mare întreg. De aceea, x trebuie să fie egal cu 1.

electric staționar în stratul de poleială. Nu este nici o îndoială că experiența fizică are un rezultat determinat. Ce putem spune însă despre problema matematică corespunzătoare, care este de cea mai mare importanță în teoria funcțiilor și în alte domenii? În conformitate cu teoria electricității, fenomenul fizic este descris de o „problemă la limită pentru o ecuație cu derivate parțiale". Aceasta este problema matematică care ne interesează pe noi; rezolvabilitatea ei este plauzibilă datorită presupusei echivalențe cu un fenomen fizic, dar prin acest raționament ea nu este cîtuși de puțin demonstrată din punct de vedere matematic. Riemann a rezolvat problema matematică în doi pași. Mai întîi, el a arătat că problema este echivalentă cu o problemă de minim: o anumită cantitate care exprimă energia curentului electric este minimizată de curentul real, în comparație cu toți ceilalți curenți posibili în condițiile date. Apoi el a enunțat ca „principiu al lui Dirichlet" faptul că o astfel de problemă de minim are soluție. Riemann nu a făcut nici cel mai mic pas spre o demonstrație matematică a celei de-a doua afirmații și acesta a fost punctul atacat de Weierstrass. Nu numai că existența minimului nu era cîtuși de puțin evidentă, dar așa cum s-a arătat, era o problemă extrem de delicată, pentru care matematica acelor timpuri nu era încă pregătită și care a fost rezolvată definitiv doar după multe decenii de cercetare intensă.

2. Exemple

Vom ilustra tipul de dificultate care intervine, cu ajutorul a două exemple.

1) Alegem două puncte A și B, aflate la o distanță d unul de altul pe o dreaptă L și cerem poligonul de perimetru minim, care pornește din A într-o direcție perpendiculară pe L, și se termină în B. Deoarece segmentul rectiliniu AB este cel mai scurt drum dintre A și B, sîntem siguri că orice drum admis în competiție are o lungime mai mare decît d, deoarece singurul drum care dă valoarea d este segmentul rectiliniu AB, care nu îndeplinește condiția impusă referitoare la direcția în A, și deci nu poate fi admis în condițiile problemei. Pe de altă parte, să considerăm drumul admisibil AOB din fig. 222. Dacă înlocuim pe O printr-un punct O', destul de apropiat de A, putem obține un drum admisibil cu o lungime care diferă oricît de puțin vrem de d. Deci dacă există un drum admisibil de lungime *minimă*, el nu poate avea o lungime care depășește pe d, și de aceea trebuie să aibă lungimea chiar egală cu d. Dar singurul drum cu această lungime nu este admisibil, așa cum am văzut mai sus. Deci nu există drum admisibil de lungime minimă și problema de minim propusă nu are soluție.

2) Așa cum se vede din fig. 223, fie C un cerc și S un punct aflat la o distanță egală cu 1 deasupra centrului său. Să considerăm clasa tuturor suprafețelor mărginite de C, care trec prin punctul S și se află deasupra lui C, astfel încît două puncte diferite să nu aibă aceeași proiecție verticală pe planul lui C. Care dintre aceste suprafețe are cea mai mică arie? Această problemă, deși pare

naturală, nu are soluție; nu există nici o suprafață admisibilă de arie minimă. Dacă nu ar fi fost impusă condiția ca suprafața să treacă prin S, soluția ar fi fost evident discul circular plan mărginit de C. Să notăm cu A aria discului. Orice altă suprafață mărginită de C trebuie să aibă o arie mai mare decît A. Însă putem găsi o suprafață admisibilă, a cărei arie depășește pe A cu oricît

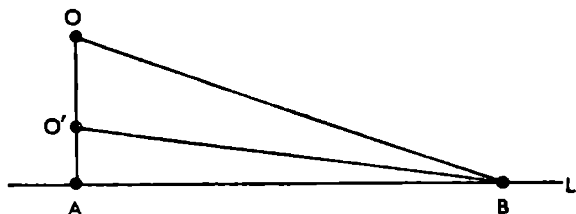


Fig. 222.

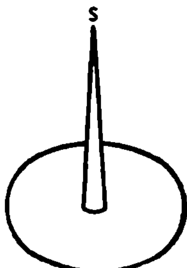


Fig. 223.

de puțin vrem. În acest scop să luăm o suprafață conică de înălțime 1, atît de subțire, încît aria ei să fie mai mică decît o margine arbitrară, pe care am dat-o de la început. Să așezăm acest con vertical pe disc cu vîrful în S, și să considerăm suprafața totală formată din suprafața conului și din acea parte a discului care se află în afara bazei conului. Se vede imediat că această suprafață, care diferă de plan doar în jurul centrului, are o arie care depășește pe A cu mai puțin decît marginea dată. Deoarece această margine poate fi aleasă oricît de mică vrem, rezultă din nou că minimul, dacă există, nu poate fi decît aria A a discului. Însă dintre toate suprafețele mărginite de C, doar discul are această arie, și deoarece discul nu trece prin S, el nu este admisibil. În consecință, problema nu are soluție.

Ne putem lipsi de exemplele mai artificiale, date de Weierstrass. Acelea pe care le-am considerat mai sus arată destul de bine că existența unui minim nu este o parte banală a unei demonstrații matematice. Să enunțăm problema în termeni mai generali și mai abstracți. Să considerăm o anumită clasă de obiecte, ca de pildă curbe sau suprafețe, fiecăreia fiindu-i atașată printr-o funcție un anumit număr, ca de pildă lungimea sau aria. Dacă există doar un număr finit de obiecte în clasa considerată, atunci evident trebuie să existe un cel mai mare și un cel mai mic printre numerele corespunzătoare. Însă dacă există o infinitate de obiecte în clasă, atunci nu este necesar să existe nici cel mai mare, nici cel mai mic număr, chiar dacă toate aceste numere sînt cuprinse între margini fixate. În general, aceste numere vor forma o mulțime infinită de puncte pe axa numerică. Să presupunem, pentru simplificare, că toate numerele sînt pozitive. Atunci, mulțimea are „o margine inferioară”, adică un punct la stînga căruia nu se află nici un număr al mulțimii și care este fie un element

al mulțimii, fie aproximat cu orice grad de precizie prin elemente ale mulțimii. Dacă α aparține mulțimii, el este cel mai mic element; în caz contrar, mulțimea nu conține un cel mai mic element. De exemplu, mulțimea de numere 1, 1/2, 1/3, 1/4, ... nu conține un cel mai mic element, deoarece marginea inferioară 0 nu aparține mulțimii. Aceste exemple ilustrează într-un mod abstract dificultățile logice legate de problema de existență. Rezolvarea matematică a unei probleme de minim nu este completă pînă ce nu am furnizat în mod explicit sau implicit o demonstrație a faptului că mulțimea valorilor asociate problemei conține un cel mai mic element.

3. Probleme elementare de extremum

În problemele elementare se cere doar o analiză atentă a noțiunilor de bază care intervin, pentru a rezolva problema existenței unei soluții. În cap. VI, § 5 a fost discutată noțiunea generală de mulțime compactă; s-a afirmat că o funcție continuă, definită pentru elementele unei mulțimi compacte, ia întotdeauna o cea mai mare și o cea mai mică valoare, undeva în mulțime. În fiecare dintre problemele elementare discutate mai înainte, valorile care intervin pot fi privite ca valori ale unei funcții de una sau mai multe variabile, care parcurg un domeniu care este fie compact, fie poate fi făcut compact, fără o modificare esențială a problemei. Într-un astfel de caz, existența unui maxim și a unui minim este asigurată. În problema lui Steiner, de exemplu, cantitatea considerată este suma a trei distanțe și ea depinde continuu de poziția punctului mobil. Deoarece domeniul acestui punct este întregul plan, nu pierdem nimic dacă includem figura într-un cerc mare și obligăm punctul să rămînă în interiorul sau pe frontiera lui. Într-adevăr, de îndată ce punctul mobil este destul de departe de cele trei puncte date, suma distanțelor sale la aceste puncte va depăși cu siguranță pe $AB + AB$, care este una dintre valorile admisibile ale funcției. Deci dacă există un minim pentru un punct constrîns să rămînă într-un cerc mare, acesta va fi, de asemenea, un minim al problemei nerestrînse. Însă este ușor de arătat că domeniul format de cerc împreună cu interiorul său este compact, deci există un minim al problemei lui Steiner.

Importanța ipotezei că domeniul variabilei independente este compact poate fi arătată prin următorul exemplu. Fiind date două curbe închise C_1 și C_2, există întotdeauna două puncte P_1 și P_2, respectiv pe C_1, C_2, care se află la cea mai mică distanță posibilă unul de altul, și două puncte Q_1, Q_2, care se află la cea mai mare distanță posibilă. Într-adevăr, distanța dintre un punct A_1 de pe C_1 și un punct A_2 de pe C_2 este o funcție continuă pe mulțimea compactă formată din toate perechile A_1, A_2 de puncte considerate. Însă, dacă cele două curbe nu sînt mărginite, ci se întind la infinit, atunci problema poate să nu aibă soluție. În cazul arătat în fig. 224 nu există nici cea mai mică și nici cea mai mare distanță între curbe; marginea inferioară a distanței este egală cu 0, în timp ce marginea superioară este infinită și nici una din ele nu

este atinsă. În unele cazuri există un minim, dar nu şi un maxim. În cazul celor două ramuri ale unei hiperbole (fig. 17), doar distanța minimă este atinsă în punctele A şi A', deoarece, evident, nu există două puncte aflate la distanță maximă.

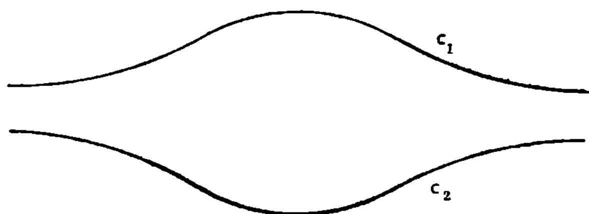


Fig. 224. Curbe între care nu există o cea mai mare sau cea mai mică distanță

Putem explica această deosebire de comportare în cazul restrîngerii artificiale a domeniului variabilelor. Să alegem un număr pozitiv arbitrar R şi să restrîngem pe x prin condiția $|x| \leqslant R$. Atunci există atît un maxim, cît şi un minim pentru fiecare din ultimele două probleme. În prima, restrîngînd frontiera în acest mod, asigurăm existența unei distanțe maxime şi a uneia minime, fiecare fiind atinsă pe frontieră. Dacă R creşte, punctele pentru care extremele sînt atinse sînt tot pe frontieră. Deci pe măsură ce R creşte, aceste puncte dispar spre infinit. În cel de-al doilea caz, distanța minimă este atinsă în interior şi oricît de mult ar creşte R, cele două puncte de minim rămîn neschimbate.

4. Dificultăți în cazuri superioare

În timp ce problema existenței nu este cîtuşi de puțin dificilă în problemele elementare în care intervin una, două sau un număr finit de variabile independente, lucrurile stau cu totul altfel în cazul principiului lui Dirichlet sau chiar cu probleme mai simple de natură asemănătoare. Cauza se află fie în faptul că domeniul variabilelor independente nu mai este compact, fie în faptul că funcția nu mai este continuă. În primul exemplu al secțiunii 2, avem un şir de drumuri $AO'B$, unde O' tinde spre punctul A. Fiecare drum al şirului satisface condițiile de admisibilitate. Însă drumurile $AO'B$ tind spre segmentul rectiliniu AB şi această limită nu mai este în mulțimea admisă. Mulțimea drumurilor admisibile este comparabilă din acest punct de vedere cu intervalul $0 < x \leqslant 1$, pentru care teorema lui Weierstrass asupra valorilor extreme nu mai este valabilă (cf. p. 330). În cel de-al doilea exemplu găsim o situație asemănătoare; dacă conurile devin din ce în ce mai subțiri, atunci şirul de suprafețe admisibile corespunzătoare vor tinde spre disc, la care se adaugă un seg-

ment de dreaptă vertical care ajunge în S. Această entitate geometrică limită nu se află însă printre suprafețele admisibile și vedem din nou că mulțimea suprafețelor admisibile nu este compactă.

Ca exemplu de dependență necontinuă putem considera lungimea unei curbe. Această lungime nu mai este funcție de un număr finit de variabile numerice,

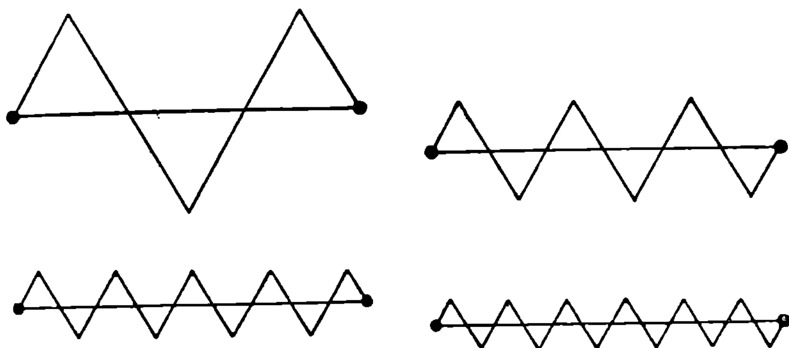


Fig. 225. Aproximarea unui segment prin linii frînte de lungime dublă

deoarece o curbă întreagă nu poate fi caracterizată printr-un număr finit de „coordonate" și lungimea nu este funcție continuă de curbă. Pentru a vedea acest lucru, să unim două puncte A și B, aflate la o distanță d unul de altul, printr-un poligon în zigzag P_n, care împreună cu segmentul AB formează n triunghiuri echilaterale. Este clar din fig. 225, că lungimea totală a lui P_n va fi exact $2d$, pentru fiecare valoare a lui n. Să considerăm acum șirul de poligoane P_1, P_2, ... Undele formate de aceste poligoane descresc în înălțime pe măsură ce numărul lor crește și este clar că poligonul P_n tinde spre dreapta AB, iar la limită perturbațiile au dispărut complet. Lungimea lui P_n este întotdeauna egală cu $2d$, oricare ar fi indicele n, în timp ce lungimea curbei limită, segmentul de dreaptă, este egală cu d. Deci lungimea nu depinde continuu de curbă.

Toate aceste exemple confirmă faptul că este necesară o anumită precauție în privința existenței unei soluții, în problemele de minim de structură mai complicată.

§ 8. PROBLEMA IZOPERIMETRICĂ

Faptul că cercul cuprinde cea mai mare arie dintre toate curbele închise de lungime dată este unul din faptele „evidente" ale matematicii, pentru care doar metodele moderne au dat o demonstrație riguroasă. Steiner a imaginat

diferite căi ingenioase de demonstrare a acestei teoreme, din care vom considera una singură.

Să începem cu ipoteza că există o soluție. Admițînd acest lucru, să presupunem că este cerută curba C cu lungimea dată L și cu aria maximă. Atunci, putem arăta cu ușurință că C trebuie să fie convexă, în sensul că segmentul de dreaptă care unește două puncte oarecare ale lui C, trebuie să se afle în întregime

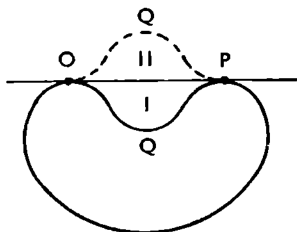


Fig. 226.

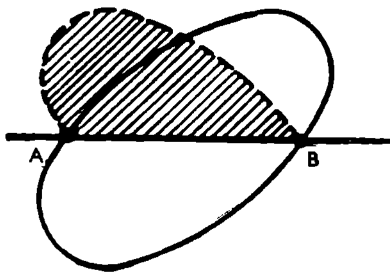


Fig. 227.

în interiorul lui C sau pe C. Într-adevăr, dacă C nu ar fi convexă ca în fig. 226, atunci ar putea fi trasat un segment, ca de pildă OP între două puncte O și P de pe C, astfel încît OP să se afle în exteriorul lui C. Arcul $OQ'P$, care este simetricul lui OQP față de dreapta OP, împreună cu arcul ORP formează o curbă de lungime L, care cuprinde o arie mai mare decît curba inițială C, deoarece ea cuprinde ariile suplimentare I și II. Aceasta contrazice ipoteza că C conține cea mai mare arie dintre toate curbele închise de lungime L. Deci C trebuie să fie convexă.

Să alegem acum două puncte A, B care împart curba soluție C în două arce de lungime egală. Atunci dreapta AB trebuie să împartă aria lui C în două părți egale, deoarece în caz contrar am putea construi simetricul față de dreapta AB (fig. 227) al suprafeței de arie mai mare, astfel încît să obținem o altă curbă de lungime L, care cuprinde o arie mai mare decît aceea a lui C. Rezultă că o jumătate din curba soluție C trebuie să rezolve următoarea problemă: să se găsească arcul de lungime $L/2$, care are extremitățile A, B pe o dreaptă și care cuprinde o arie maximă între el și segmentul de dreaptă. Vom arăta acum că soluția acestei noi probleme este un semicerc, astfel încît întreaga curbă C, care rezolvă problema izoperimetrică, este un cerc. Să presupunem că arcul AOB rezolvă noua problemă. Este suficient să arătăm că orice unghi înscris în fig. 228, ca de pildă $\sphericalangle AOB$, este un unghi drept, pentru că acest lucru va demonstra că AOB este un semicerc. Să presupunem contrariul, și anume că unghiul AOB nu este de 90°. Atunci putem înlocui fig. 228 prin alta, fig. 229, în care ariile hașurate și lungimea arcului AOB nu sînt modificate, în timp ce aria triunghiulară este mărită, făcînd $\sphericalangle AOB$ egal sau cel puțin apropiat de

90°. Astfel, fig. 229 dă o arie mai mare decît cea inițială (cf. p. 348). Am pornit însă de la ipoteza că fig. 228 rezolvă problema, astfel încît fig. 229 nu poate da o arie mai mare. Această contradicție arată că pentru orice punct O, $\sphericalangle AOB$ trebuie să fie un unghi drept, ceea ce completează demonstrația.

Proprietatea izoperimetrică a cercului poate fi exprimată sub forma unei inegalități. Dacă L este circumferința cercului, aria lui este $L^2/4\pi$ și de aceea

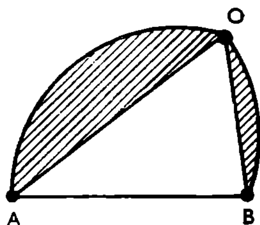


Fig. 228.

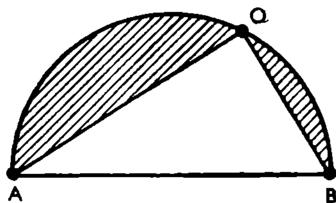


Fig. 229.

trebuie să avem inegalitatea izoperimetrică $A \leqslant L^2/4\pi$ între aria A și lungimea L a oricărei curbe închise, semnul egal avînd loc numai pentru cerc.

*După cum reiese din discuția cuprinsă în § 7, demonstrația lui Steiner are numai o valoare condițională: „dacă există o curbă de lungime L cu arie maximă, atunci ea trebuie să fie un cerc". Pentru a demonstra premisa, este necesar un raționament esențial nou. Mai întîi demonstrăm o teoremă elementară, referitoare la poligoanele închise P_n, cu un număr par $2n$ de laturi: dintre toate aceste $2n$-goane de aceeași lungime $2n$-gonul regulat are aria maximă. Demonstrația merge pe drumul raționamentului lui Steiner, cu următoarele modificări. Aici nu există nici o dificultate în legătură cu problema existenței, deoarece un $2n$-gon, împreună cu lungimea și aria lui, depinde continuu de cele $4n$ coordonate ale vîrfurilor lui, care fără a restrînge generalitatea, pot fi obligate să rămînă într-o mulțime compactă de puncte din spațiul $4n$-dimensional. În consecință, în această problemă pentru poligoane putem începe fără grijă cu ipoteza că un anumit poligon P *este* soluția problemei, și pe această bază putem analiza proprietățile lui P. Ca și în demonstrația lui Steiner, rezultă că P trebuie să fie convex. Să demonstrăm acum că toate cele $2n$ muchii ale lui P trebuie să aibă aceeași lungime. Într-adevăr, să presupunem că două muchii vecine AB și BC au lungimi diferite; atunci, am putea secționa din P triunghiul ABC și l-am putea înlocui printr-un triunghi isoscel $AB'C$, în care $AB' + B'C = AB + BC$ și care are o arie mai mare (cf. § 1). Astfel, am obține un poligon P' cu același perimetru și cu o arie mai mare, contrar ipotezei că P era poligonul optim cu $2n$ laturi. De aceea, toate laturile lui P trebuie să aibă aceeași lungime și rămîne de arătat numai că P este regulat; pentru aceasta, este de ajuns să știm că toate vîrfurile lui P se află pe un cerc. Raționamentul urmează calea lui Steiner. Mai întîi arătăm că orice diagonală care

unește două vîrfuri opuse, de pildă primul cu cel de-al $(n+1)$-lea, împarte aria în două părți egale. Apoi demonstrăm că toate vîrfurile uneia dintre aceste părți se află pe un semicerc. Amănuntele, care urmează cu exactitate calea precedentă, sînt lăsate ca exercițiu pe seama cititorului.

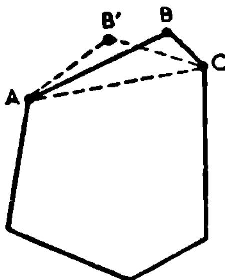


Fig. 230.

Existența, împreună cu soluția problemei izoperimetrice, pot fi obținute acum printr-un proces de trecere la limită, în care numărul vîrfurilor tinde spre infinit, iar poligonul regulat optim tinde spre un cerc.

Raționamentul lui Steiner nu este potrivit pentru a demonstra proprietatea izoperimetrică corespunzătoare a sferei, în spațiul cu trei dimensiuni. O tratare oarecum diferită și mai complicată a fost dată de Steiner și ea operează în trei dimensiuni tot atît de bine ca și în două, dar, deoarece nu poate fi adaptată atît de repede pentru a da demonstrația existenței, o vom omite. De fapt, demonstrarea proprietății izoperimetrice a sferei este un lucru mult mai dificil decît pentru cazul cercului; într-adevăr, o demonstrație completă și riguroasă a fost dată pentru prima dată mult mai tîrziu, într-un articol destul de dificil al lui H. A. Schwarz. Proprietatea izoperimetrică tridimensională poate fi exprimată prin inegalitatea

$$36\pi V^2 \leqslant A^3$$

dintre aria suprafeței A și volumul V al oricărui corp tridimensional închis, egalitatea avînd loc numai pentru sferă.

* § 9. PROBLEME DE EXTREMUM CU CONDIȚII LA LIMITĂ. LEGĂTURA DINTRE PROBLEMA LUI STEINER ȘI PROBLEMA IZOPERIMETRICĂ

Rezultate interesante apar în probleme de extremum, atunci cînd domeniul variabilei este restrîns prin condiții la limită. Teorema lui Weierstrass, în virtutea căreia într-un domeniu compact o funcție continuă își atinge o cea mai

mare şi o cea mai mică valoare, nu exclude posibilitatea ca valorile extreme să fie atinse pe frontiera domeniului. Un exemplu simplu, aproape banal, este dat de funcţia $u = x$. Dacă variabila x nu este restrînsă şi poate varia de la $-\infty$ la $+\infty$, atunci domeniul B al variabilei independente este întreaga axă numerică şi deci este uşor de înţeles că funcţia $u = x$ nu are o cea mai mare sau o cea mai mică valoare. Dacă însă domeniul B este limitat de o frontieră, de pildă $0 \leqslant x \leqslant 1$, atunci există o cea mai mare valoare 1, atinsă în extremitatea din dreapta, şi o cea mai mică valoare 0, atinsă în extremitatea din stînga. Însă aceste valori extreme nu sînt reprezentate printr-un vîrf sau o depresiune a curbei care reprezintă funcţia; ele nu sînt extreme faţă de o vecinătate completă, care se întinde de ambele părţi ale punctului. Ele se modifică îndată ce extindem intervalul, pentru că ele rămîn la extremităţi. Pentru un vîrf, sau o depresiune adevărată a unei funcţii, caracterul extrem se referă întotdeauna la o vecinătate plină a punctului în care este atinsă valoarea; ea nu este influenţată de mici modificări ale frontierei. Un astfel de extremum persistă, chiar după o variaţie liberă a variabilei independente în domeniul B, cel puţin într-o vecinătate suficient de mică. Distincţia dintre astfel de extreme „libere" şi acelea atinse la frontieră clarifică multe situaţii, care aparent sînt foarte diferite. Pentru funcţiile de o variabilă, desigur, distincţia este doar aceea dintre funcţiile monotone şi cele nemonotone şi deci nu duce la observaţii deosebit de interesante. Există însă multe exemple însemnate de extreme atinse la frontiera domeniului de variaţie de către funcţii de mai multe variabile.

Aceasta se poate întîmpla, de pildă, în problema triunghiului a lui Schwarz. Domeniul de variaţie al celor trei variabile independente constă din toate tripletele de puncte, cîte unul pe fiecare dintre cele trei laturi ale triunghiului ABC. Soluţia problemei implică două alternative: fie minimul este atins cînd toate cele trei puncte, variabile independente P, Q, R, se află în interiorul laturilor corespunzătoare ale triunghiului, în care caz minimul este dat de triunghiul ortic sau minimul este atins în poziţia frontieră. În acest ultim caz, două din punctele P, Q, R coincid cu extremitatea comună a intervalelor corespunzătoare, iar „triunghiul" înscris de perimetru minim este înălţimea coborîtă din acest vîrf, socotită de două ori. Astfel natura soluţiei este cu totul diferită, după alternativele care au loc.

În problema lui Steiner a celor trei localităţi, domeniul de variaţie al punctului P este întregul plan, în care cele trei puncte date A, B, C pot fi considerate ca puncte frontieră. Şi aici există două alternative, care duc la două tipuri cu totul diferite de soluţii: sau minimul este atins în interiorul triunghiului ABC, care este cazul celor trei unghiuri egale, sau este atins într-un punct frontieră C. O pereche asemănătoare de alternative există pentru problema complementară.

Ca un ultim exemplu putem considera problema izoperimetrică, modificată prin condiţii la limită restrictive. Vom obţine în acest mod o legătură sur-

prinzătoare între problema izoperimetrică şi problema lui Steiner, şi în acelaşi timp poate, cel mai simplu exemplu al unui nou tip de problemă de extremum. În problema inițială, variabila independentă, o curbă închisă de lungime dată, poate fi modificată în mod arbitrar față de forma circulară şi orice curbă deformată de acest fel este o curbă admisibilă, astfel încît avem un minim liber adevărat. Să considerăm acum următoarea problemă modificată: curbele C

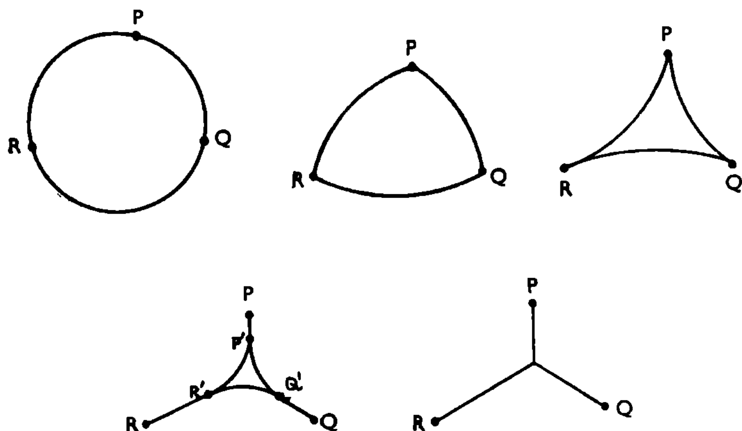


Fig. 231—235. Figuri izoperimetrice care tind spre soluția problemei lui Steiner

considerate trebuie să includă în interiorul lor sau să treacă prin trei puncte date P, Q, R, aria A este dată şi lungimea L trebuie minimalizată. Aceasta reprezintă o condiție la limită.

Este clar că dacă A este destul de mare, cele trei puncte P, Q, R nu vor influența problema. Ori de cîte ori cercul circumscris triunghiului PQR are o arie mai mică sau egală cu A, soluția va fi un cerc de arie A, care cuprinde cele trei puncte. Dar ce se întîmplă dacă A este mai mic? Dăm aici răspunsul, însă omitem amănuntele demonstrației, cu toate că ea nu ne-ar fi inaccesibilă. Să caracterizăm soluțiile pentru un şir de valori ale lui A, care descreşte spre 0. Îndată ce A devine mai mic decît aria cercului circumscris, cercul izoperimetric inițial se descompune în trei arce, toate de aceeaşi rază, care formează un triunghi circular convex, cu vîrfurile P, Q, R (fig. 232). Acest triunghi este soluția problemei; dimensiunile lui pot fi determinate cunoscînd valoarea dată a lui A. Dacă A descreşte în continuare, razele acestor arce vor creşte şi arcele vor deveni din ce în ce mai drepte, pînă ce A este egal cu aria triunghiului PQR, cînd soluția este chiar acest triunghi. Dacă A devine acum mai mic, atunci soluția va consta din nou din trei arce circulare, care au aceeaşi rază şi formează un triunghi cu vîrfurile în P, Q, R. De această dată însă, triunghiul

stee concav şi arcele se află în interiorul triunghiului PQR (fig. 233). Pe măsură ce A descreşte în continuare, va veni un moment cînd, pentru o anumită valoare a lui A, două din arcele concave devin tangente într-un vîrf R. La o descreştere ulterioară a lui A, nu mai este posibil să construim un triunghi circular de tipul precedent. Apare un fenomen nou. Soluţia este dată tot de un triunghi circular concav, însă unul dintre vîrfurile lui R' s-a detaşat de vîrful corespunzător R şi soluţia constă acum dintr-un triunghi circular PQR' plus segmentul de dreaptă RR', socotit de două ori (pentru că el este parcurs de la R' la R şi înapoi). Acest segment de dreaptă este tangent celor două arce tangente în R'. Dacă A descreşte în continuare, procesul de separare va avea loc şi în celelalte vîrfuri. La un anumit moment, obţinem ca soluţie un triunghi circular, format din trei arce de raze egale, tangente două cîte două şi care formează un triunghi circular echilateral $P'Q'R'$, la care se mai adaugă trei segmente de dreaptă $P'P$, $Q'Q$, $R'R$, socotite de două ori (fig. 234). Dacă în cele din urmă A se anulează, atunci triunghiul circular se reduce la un punct şi revenim la soluţia problemei lui Steiner; aceasta este, prin urmare, un caz limită al problemei izoperimetrice modificate.

Dacă P, Q, R formează un triunghi obtuz, cu un unghi mai mare de 120°, atunci procesul de contractare duce la soluţia corespunzătoare a problemei lui Steiner, pentru că atunci arcele circulare se contractă spre vîrful obtuz. Soluţiile problemei generalizate a lui Steiner (fig. 216—218) pot fi obţinute prin procese de trecere la limită de natură similară.

§ 10. CALCULUL VARIAŢIONAL

1. Introducere

Problema izoperimetrică reprezintă, poate, cel mai vechi exemplu dintr-o clasă largă de probleme importante, asupra cărora a atras atenţia Jean Bernoulli în 1696. În „Acta Eruditorum”, marea revistă ştiinţifică a timpului, el a pus următoarea problemă „brahistocronă”. Imaginaţi-vă o particulă constrînsă să alunece fără frecare în lungul unei curbe care uneşte un punct A cu un punct mai coborît B. Dacă particula poate cădea doar sub influenţa gravitaţiei, în lungul cărei curbe timpul de coborîre va fi minim? Este uşor de văzut că particula aflată în cădere va avea nevoie de intervale diferite de timp pentru drumuri diferite. Dreapta nu asigură în nici un caz cel mai scurt timp; acelaşi lucru se poate spune despre arcul de cerc şi despre oricare altă curbă elementară. Bernoulli s-a lăudat că posedă o soluţie minunată, pe care el nu vrea s-o publice imediat, pentru a provoca pe cei mai mari matematicieni ai timpului să-şi încerce măiestria cu acest nou tip de problemă matematică. Printre alţii, el a provocat pe fratele său mai mare, Jacques, cu care era, în acel timp, într-o ceartă înverşunată şi pe care l-a descris în mod public ca incompetent, să rezolve problema. Matematicienii şi-au dat seama imediat de natura

diferită a problemei brahistocrone. În problemele cercetate pînă atunci cu ajutorul calculului diferențial, cantitatea minimizată depindea numai de una sau mai multe variabile numerice; în această problemă, cantitatea considerată, timpul de coborîre, depinde de întreaga curbă, și aceasta este o deosebire esențială, care scoate problema din domeniul calculului diferențial, sau al oricărei alte metode cunoscute pe atunci.

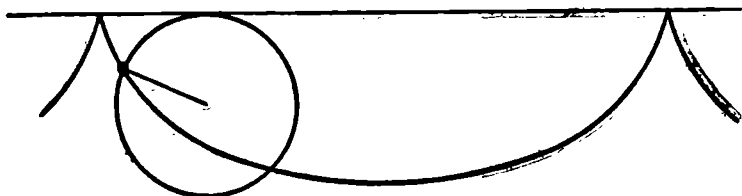


Fig. 236. Cicloida

Noutatea problemei — după cît se pare, proprietatea izoperimetrică a cercului nu a fost recunoscută în mod clar ca fiind de aceeași natură — a fascinat pe matematicienii contemporani, cu atît mai mult atunci cînd s-a aflat că soluția este dată de cicloidă, o curbă care tocmai atunci fusese descoperită. (Reamintim definiția cicloidei: locul geometric al unui punct de pe circumferința unui cerc, care se rostogolește fără alunecare în lungul unei drepte, așa cum se arată în fig. 236.) Această curbă a fost pusă în legătură cu probleme interesante de mecanică, în special cu construirea unui pendul ideal. Huygens descoperise că un punct material ideal, care oscilează fără frecare sub influența gravitației în lungul unei cicloide verticale, are o perioadă de oscilație independentă de amplitudinea mișcării. Pe un drum circular care este descris de un pendul obișnuit, această independență este adevărată doar în mod aproximativ și acest fapt a fost considerat ca un neajuns al folosirii pendulului pentru ceasornicele de precizie. Cicloida a fost onorată cu numele de tautocronă; acum, ea a dobîndit noul titlu de brahistocronă.

2. Calculul variațional. Principiul lui Fermat din optică

Din diferitele moduri în care frații Bernoulli și alți matematicieni au rezolvat problema brahistocronei, vom expune îndată unul dintre cele mai originale. Primele metode au un caracter mai mult sau mai puțin particular, adaptat acestei probleme speciale. Dar foarte curînd, Euler și Lagrange (1736—1813) au dezvoltat metode mai generale pentru rezolvarea problemelor de extremum, în care elementul independent nu era o singură variabilă numerică sau un număr finit de variabile de acest fel, ci o curbă, sau funcție, sau chiar un sistem de curbe (funcții). Noua metodă de rezolvare a acestor probleme a fost numită *calculul variațional.*

Nu putem descrie aici aspectele tehnice ale acestei ramuri a matematicii sau să pătrundem mai adînc în discutarea diferitelor probleme. Calculul variațional are numeroase aplicații în fizica teoretică. S-a observat de mult că fenomenele naturale satisfac adesea anumite condiții de maxim sau de minim. După cum am văzut, Heron din Alexandria și-a dat seama că reflexia unei

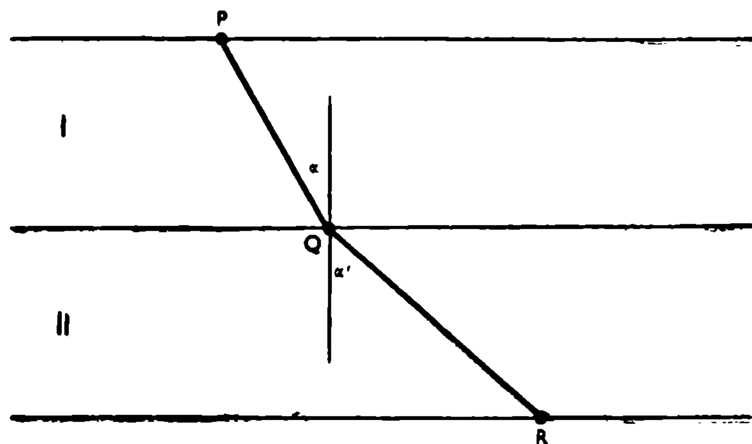


Fig. 237. Refracția unei raze de lumină

raze de lumină într-o oglindă plană poate fi descrisă printr-un principiu de minim. Fermat, în secolul al XVII-lea, a făcut pasul următor, observînd că și legea refracției luminii poate fi formulată în limbajul unui principiu de minim. Este binecunoscut faptul că drumul parcurs de o rază de lumină, care trece dintr-un mediu omogen în altul, este frînt în punctul de trecere. Astfel, în fig. 237, o rază de lumină, care pornește din punctul P, aflat în mediul superior în care viteza este v și ajunge în punctul R, aflat în mediul inferior în care viteza este w, va parcurge un drum PQR. Legea empirică găsită de Snell (1591—1626) afirmă că drumul constă din două segmente rectilinii PQ și QR, care formează unghiuri α, α' cu normala, determinate de condiția $\sin\alpha/\sin\alpha' = v/w$. Cu ajutorul calculului diferențial, Fermat a demonstrat că acest drum este astfel, încît timpul necesar razei de lumină pentru a porni din P și a ajunge în R este minim, adică este mai mic decît acela necesar pentru a parcurge oricare alt drum dintre P și P'. Așadar, după 1600 de ani, legea reflexiei luminii a lui Heron a fost completată cu legea tot atît de importantă și asemănătoare a refracției.

Fermat a generalizat enunțul acestei legi, astfel încît să includă cazul unor suprafețe curbe de discontinuitate dintre medii, ca de pildă suprafețele sferice folosite la lentile. În acest caz, rămîne în vigoare afirmația că lumina

parcurge un drum în lungul căruia timpul necesar este minim, în comparație cu timpul necesar luminii pentru a descrie orice alt drum care unește cele două puncte. În sfîrșit, Fermat a considerat un sistem optic arbitrar, în care viteza luminii variază într-un mod cunoscut, de la punct la punct, așa cum se întîmplă în atmosferă. El a împărțit mediul continuu neomogen în straturi subțiri, în fiecare dintre acestea viteza luminii fiind aproximativ constantă și a imaginat acest mediu ca fiind înlocuit cu altul, în care viteza este într-adevăr constantă în fiecare strat. Atunci el a putut aplica din nou principiul său, trecînd de la fiecare strat la cel următor. Făcînd ca grosimea straturilor să tindă spre zero, el a ajuns la *principiul general al lui Fermat din optica geometrică*: într-un mediu neomogen, o rază de lumină care trece de la un punct la altul urmează un drum, pentru care timpul necesar este minim, în comparație cu toate drumurile care unesc cele două puncte. Acest principiu a fost de cea mai mare importanță, nu numai din punct de vedere teoretic, dar și în optica geometrică practică. Tehnica calculului variațional aplicată acestui principiu constituie baza pentru calculul sistemelor de lentile.

Principiile de minim au devenit dominante și în alte ramuri ale fizicii. S-a observat că echilibrul stabil al unui sistem mecanic este atins dacă sistemul este astfel dispus încît „energia sa potențială" să fie minimă. Ca exemplu, să considerăm un lanț omogen flexibil, suspendat de cele două capete și aflat sub influența gravitației. Atunci lanțul va lua o formă în care energia sa potențială este minimă. În acest caz, energia potențială este determinată de înălțimea centrului de greutate față de o axă fixă. Curba în lungul căreia atîrnă lanțul se numește *lănțișor* și forma sa exterioară ne amintește întrucîtva parabola.

Nu numai legile de echilibru, dar și acelea de mișcare sînt dominate de principii de maxim și de minim. Euler a fost acela care a obținut primele idei clare despre aceste principii, în timp ce persoane înclinate spre meditații speculative filozofice sau mistice, ca de pildă Maupertuis (1698—1759), nu au fost capabile să deosebească formulările exacte matematice de ideile nebuloase referitoare la „intenția lui Dumnezeu de a conduce fenomenele fizice printr-un principiu general de perfecțiune maximă". Principiile variaționale ale lui Euler din fizică, redescoperite și extinse de matematicianul irlandez W. R. Hamilton (1805—1865), s-au arătat a fi printre cele mai puternice instrumente ale mecanicii, opticii și electrodinamicii, avînd numeroase aplicații în tehnică. Cercetări recente din fizica teoretică — teoria relativității și teoria cuantică — sînt exemple care dovedesc pe deplin însemnătatea metodelor calculului variațional.

3. Tratarea lui Bernoulli a problemei brahistocronei

Prima metodă dezvoltată pentru rezolvarea problemei brahistocronei de către Jacques Bernoulli poate fi înțeleasă cu cunoștințe tehnice comparativ reduse. Începem cu faptul cunoscut din mecanică, că un punct material care

cade în lungul curbei, începînd din punctul A, în care inițial are viteza nulă, va avea, în orice punct P, o viteză proporțională cu $\sqrt{h}$, unde h este distanța verticală de la A la P, adică $v = c\sqrt{h}$, unde c este o constantă. Acum să înlocuim problema dată cu una puțin deosebită. Împărțim spațiul în mai multe straturi orizontale subțiri, fiecare avînd grosimea d, și presupunem pentru

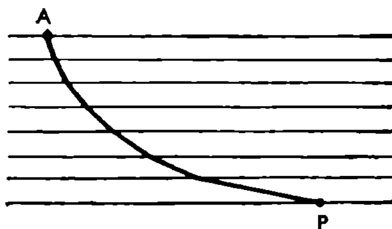


Fig. 238.

moment că viteza particulei mobile nu variază continuu, ci face mici salturi de la strat la strat, astfel încît în primul strat adiacent lui A viteza este $c\sqrt{d}$, în al doilea $c\sqrt{2d}$, iar în cel de-al n-lea strat, ea este $c\sqrt{nd} = c\sqrt{h}$, unde h este distanța verticală de la A la P (fig. 238). Dacă considerăm această problemă, atunci există numai un număr finit de variabile. În fiecare strat, drumul trebuie să fie un segment de dreaptă și nu se pune nici o problemă de existență, iar soluția trebuie să fie un poligon. Singura problemă care rămîne este determinarea vîrfurilor sale. Conform principiului de minim pentru legea refracției simple, în fiecare pereche de straturi succesive, mișcarea de la P la R prin Q trebuie să fie astfel încît, atunci cînd P și R sînt fixate, Q trebuie să dea cel mai scurt timp posibil. Deci trebuie să se aplice următoarea „lege de refracție":

$$\frac{\sin \alpha}{\sqrt{nd}} = \frac{\sin \alpha'}{\sqrt{(n+1)d}}.$$

Aplicarea repetată a acestui raționament ne conduce la șirul de egalități

$$(1) \qquad \frac{\sin \alpha_1}{\sqrt{d}} = \frac{\sin \alpha_2}{\sqrt{2d}} = \dots,$$

unde α_n este unghiul dintre latura aflată în cel de-al n-lea strat și verticală.

Apoi Bernoulli își imaginează că grosimea d devine din ce în ce mai mică, tinzînd spre zero, astfel încît poligonul obținut ca soluție a problemei aproximative tinde spre soluția dorită a problemei inițiale. În această trecere la limită, egalitățile (1) rămîn în vigoare și, de aceea, Bernoulli, trage concluzia că soluția trebuie să fie o curbă C, care are următoarea proprietate: dacă α este

unghiul făcut de tangentă cu verticala într-un punct oarecare P al lui C, iar h este distanța verticală a lui P la dreapta orizontală care trece prin A, atunci $\sin \alpha/\sqrt{h}$ este constant pentru toate punctele P ale lui C. Se poate arăta foarte ușor că această proprietate caracterizează cicloida.

,,Demonstrația'' lui Bernoulli este un exemplu tipic de raționament matematic ingenios și valoros care, în același timp, nu poate fi considerat cîtuși de puțin riguros. În el există mai multe ipoteze tacite și justificarea lor ar fi mai complicată și mai lungă decît raționamentul însuși. De exemplu, au fost presupuse existența unei soluții C și faptul că soluția problemei aproximative aproximează soluția reală. Problema valorii intrinseci a considerațiilor euristice de acest tip merită, desigur, a fi discutată, însă ne-ar duce mult prea departe.

4. Geodezice pe o sferă. Geodezicele și maxi-minimele

În introducerea la acest capitol am menționat problema găsirii celor mai scurte arce care unesc două puncte date pe o suprafață. Pe o sferă, după cum se arată în geometria elementară, aceste ,,geodezice'' sînt arce de cercuri mari. Fie P și Q două puncte (nu diametral opuse) de pe o sferă și fie c cel mai scurt arc de cerc mare care trece prin P și Q. Atunci se pune problema ce reprezintă arcul mai lung c' al aceluiași cerc mare? Desigur, el nu dă lungimea minimă și nici nu poate da lungimea maximă a curbelor care unesc pe P cu Q, deoarece

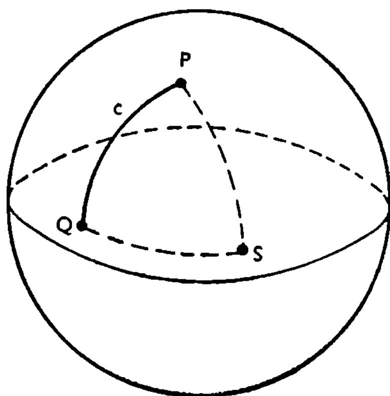


Fig. 239. Geodezice pe o sferă

pot fi trasate curbe între P și Q, de lungime arbitrară. Răspunsul este că c' rezolvă o problemă de maxi-minim. Să considerăm un punct arbitrar S aflat pe un cerc mare al sferei; se cere să se găsească cel mai scurt drum cuprins între P și Q care trece prin S. Desigur, minimul este dat de o curbă care

constă din două arce mici ale cercurilor mari care trec prin P, S şi Q, S. Să căutăm acum o poziţie a punctului S, pentru care această distanţă minimă să devină cît se poate de mare.[1] Soluţia este următoarea: S trebuie să fie astfel încît PSQ să fie arcul mai lung c' al cercului mare care trece prin punctele P şi Q. Putem modifica problema, căutînd mai întîi drumul de lungime minimă de la P la Q, care trece prin n puncte date S_1, S_2, ..., S_n de pe sferă şi căutînd apoi să determinăm punctele $S_1, \ldots, S_n$, astfel încît această lungime minimă să devină cît se poate de mare. Soluţia este dată de un drum aflat pe cercul mare care uneşte pe P cu Q, însă acest drum înconjură sfera de atîtea ori de cîte ori el trece prin punctele diametral opuse punctelor P şi Q, adică de exact n ori.

Acest exemplu de problemă de maxi-minim este tipic pentru o clasă largă de probleme ale calculului variaţional, care au fost studiate cu un mare succes prin metodele dezvoltate de Morse şi de alţii.

§ 11. SOLUŢII EXPERIMENTALE ALE UNOR PROBLEME DE MINIM. EXPERIENŢE CU PELICULE DE SĂPUN

1. Introducere

De obicei este foarte greu, iar uneori chiar imposibil, să rezolvăm problemele variaţionale explicit, cu ajutorul unor formule sau construcţii geometrice care folosesc elemente simple, cunoscute. În schimb, adesea ne mulţumim să demonstrăm doar existenţa unei soluţii în anumite condiţii şi apoi să cercetăm proprietăţile soluţiei. În multe cazuri în care o astfel de demonstraţie de existenţă pare a fi mai mult sau mai puţin dificilă, este stimulatoare realizarea condiţiilor matematice ale problemei prin artificii fizice corespunzătoare sau mai degrabă considerarea problemei matematice ca interpretare a unui fenomen fizic. Existenţa fenomenului fizic reprezintă atunci soluţia problemei matematice. Desigur, aceasta este doar o consideraţie de plauzibilitate şi nu o demonstraţie matematică, deoarece rămîne problema dacă interpretarea matematică a evenimentului fizic este adecvată în mod strict sau dacă ea dă numai o imagine inadecvată a realităţii fizice. Uneori astfel de experienţe, chiar dacă sînt efectuate doar mintal, sînt convingătoare chiar şi pentru matematicieni. În secolul al XIX-lea, multe teoreme fundamentale din teoria funcţiilor au fost descoperite de Riemann, imaginîndu-şi experienţe simple referitoare la scurgerea electricităţii în foi metalice subţiri.

În acest paragraf dorim să discutăm, pe baza demonstraţiilor experimentale, una dintre cele mai profunde teoreme ale calculului variaţional. Această problemă a fost numită problema lui Plateau, deoarece Plateau (1801—1883), cunoscut fizician de origine belgiană, a făcut experienţe interesante referitoare

la acest subiect. Problema însăși este mult mai veche și își are originea în fazele inițiale ale calculului variațional. Sub forma cea mai simplă, ea se enunță în felul următor: să se găsească suprafața de arie minimă mărginită de un contur închis, dat, din spațiu. Vom mai discuta și experiențe legate de cîteva probleme înrudite și va rezulta că în acest mod pot fi clarificate cîteva din rezultatele noastre precedente, ca și anumite probleme matematice de un tip nou.

2. Experiențe cu pelicule de săpun

Din punct de vedere matematic, problema lui Plateau conduce la rezolvarea unei „ecuații cu derivate parțiale" sau a unui sistem de astfel de ecuații. Euler a arătat că toate suprafețele minime (neplane) trebuie să aibă formă de șa și că curbura medie[2] trebuie să fie nulă în fiecare punct. S-a arătat în secolul trecut că soluția există în numeroase cazuri particulare, însă existența soluției în cazul general a fost demonstrată abia recent de J. Douglas și de T. Radó.

Experiențele lui Plateau dau imediat soluțiile fizice pentru contururi foarte generale. Dacă scufundăm un contur închis, confecționat din sîrmă, într-un lichid cu tensiune superficială mică, retrăgîndu-l apoi, se formează o peliculă de forma unei suprafețe minima, de cea mai mică arie, subîntinsă de contur. (Presupunem că putem neglija gravitația și alte forțe care interferează cu ten-

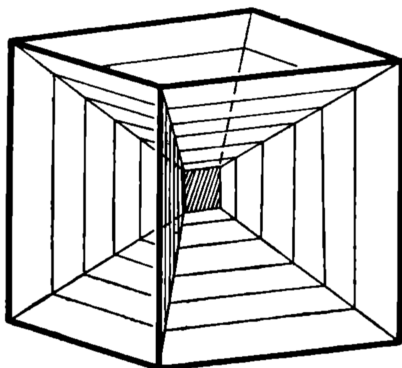


Fig. 240. Un contur cubic care subîntinde un sistem de pelicule de săpun format din 13 suprafețe aproape plane

[2] Curbura medie a unei suprafețe într-un punct P se definește în felul următor: să considerăm normala la suprafață în punctul P și toate planele care o conțin. Aceste plane vor intersecta suprafața după curbe, care în general au diferite curburi în punctul P. Să considerăm acum curbele de curbură minimă, respectiv maximă. (În general, planele care conțin aceste curbe vor fi perpendiculare.) Semisuma acestor două curburi este curbura medie a suprafeței în punctul P.

dința peliculei de a lua o poziție de echilibru stabil prin atingerea ariei minime, și deci cea mai mică valoare posibilă a energiei potențiale corespunzătoare tensiunii superficiale.) O rețetă bună pentru prepararea unui astfel de lichid este următoarea: dizolvați 10 g de oleat de sodiu uscat, pur, în 500 g apă distilată și amestecați 15 cm^3 de soluție cu 11 cm^3 de glicerină. Peliculele obținute cu această soluție și cu contururi confecționate din sîrmă de alamă sînt relativ stabile. Contururile nu trebuie să depășească 12—15 cm în diametru.

Cu această metodă este foarte ușor de „rezolvat" problema lui Plateau, confecționînd, pur și simplu, un contur de forma dorită. Se obțin modele frumoase cu ajutorul unor contururi poligonale, confecționate din sîrmă, care constituie un șir de muchii ale unui poliedru regulat. În particular, este interesant de scufundat conturul unui cub într-o astfel de soluție. Mai întîi, rezultatul este un sistem format din diferite suprafețe, care se intersectează sub unghiuri de 120°. (Dacă retragem cu grijă cubul, vom obține 13 suprafețe aproape plane.) Apoi putem perfora și distruge unele dintre aceste suprafețe, pentru a obține o singură suprafață, mărginită de un poligon închis. În acest mod pot fi formate mai multe suprafețe frumoase. Aceeași experiență poate fi efectuată și cu ajutorul unui tetraedru.

3. Noi experiențe cu problema lui Platean

Cadrul experiențelor cu pelicule de săpun și cu suprafețe minima este mai vast decît aceste demonstrații originale, făcute de Plateau. În ultimii ani, problema suprafețelor minima a fost studiată în cazuri în care este dat nu un singur contur, ci sînt date mai multe contururi și în care structura topologică a suprafeței este mai complicată. De exemplu, suprafața ar putea fi unilateră sau de gen diferit de zero. Aceste probleme mai generale produc o varietate uimitoare de fenomene geometrice, care pot fi exemplificate cu ajutorul experiențelor făcute cu pelicule de săpun. În legătură cu acest lucru, este foarte util ca contururile de sîrmă să fie flexibile și să studiem efectul deformărilor frontierelor date asupra soluției.

Vom descrie mai multe exemple:

1) Dacă conturul este un cerc, obținem un disc circular plan. Dacă deformăm continuu cercul frontieră, ne-am putea aștepta ca suprafața minimă să-și păstreze structura topologică a discului. Dar nu se întîmplă acest lucru. Dacă deformăm frontiera pînă ce ea capătă forma indicată în fig. 241, obținem o suprafață minima, care nu mai este simplu conexă ca discul, ci este o bandă a lui Moebius unilateră. Reciproc, am putea începe cu acest contur și cu o peliculă de săpun de forma unei benzi a lui Moebius. Putem deforma conturul de sîrmă trăgînd de mînere sudate de el (fig. 241). În acest proces vom ajunge la un moment în care structura topologică a peliculei se modifică brusc, astfel încît suprafața are din nou tipul unui disc simplu conex (fig. 242). Inversînd deformarea, obținem din nou banda lui Moebius. În acest proces al-

ternativ de deformare, transformarea suprafeței simplu conexe în banda lui Moebius are loc într-un stadiu mai avansat. Aceasta arată că trebuie să existe un șir de forme ale conturului, pentru care atît banda lui Moebius, cît și suprafața simplu conexă sînt stabile, adică furnizează minime relative. Însă atunci cînd banda lui Moebius are o arie mult mai mică decît cealaltă suprafață, aceasta din urmă este prea instabilă pentru a putea fi formată.

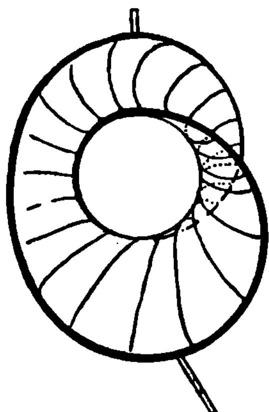


Fig. 241. Suprafață unilateră (banda lui Moebius)

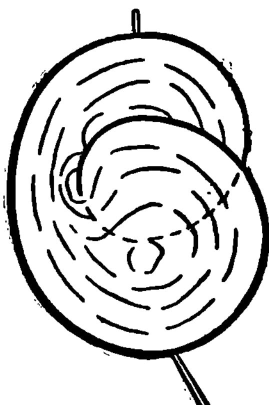


Fig. 242. Suprafață bilateră

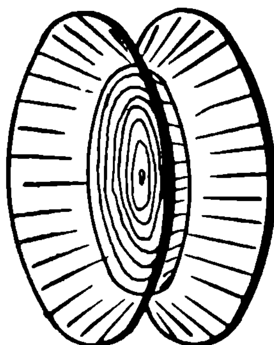
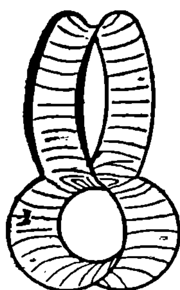


Fig. 243. Un sistem de trei suprafețe

2) Putem întinde o suprafață minima de rotație între două cercuri. După extragerea contururilor de sîrmă din soluție nu găsim o suprafață simplă, ci un sistem format din trei suprafețe care se intersectează după unghiuri de 120°, dintre care una este un disc circular simplu, paralel cu cercurile frontieră date (fig. 243). Distrugînd această suprafață intermediară, obținem o

suprafață catenoidă clasică (catenoida este suprafața obținută prin rotirea lănțișorului de la p. 401, în jurul unei drepte perpendiculare la axa sa de simetrie). Dacă cele două cercuri frontieră sînt îndepărtate, la un anumit moment catenoida dublu conexă minima devine nestabilă. În acest moment, catenoida se separă în două discuri. Acest proces, desigur, nu este reversibil.

3) Un alt exemplu semnificativ este dat de conturul din fig. 244—246, în care pot fi subîntinse trei suprafețe minima diferite. Fiecare este mărginită



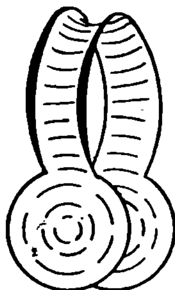


Fig. 244—245. Contur care subîntinde trei suprafețe diferite de genul 0 și 1

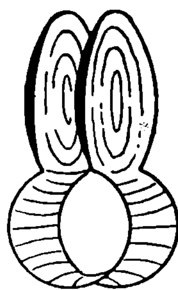


Fig. 246.

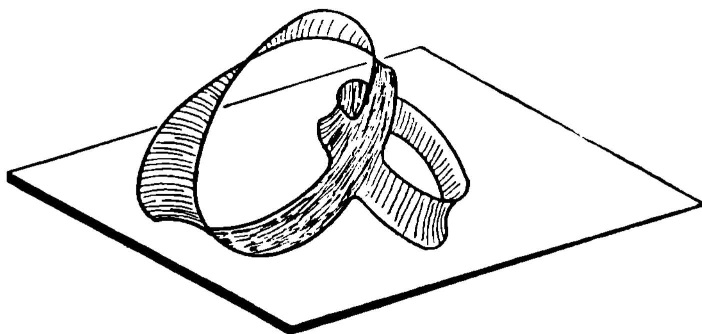


Fig. 247. Suprafață minimă unilateră, de structură topologică superioară, cu un singur contur

de aceeași curbă închisă simplă; una (fig. 244) are genul 1, în timp ce celelalte două sînt simplu conexe și oarecum simetrice una față de alta. Ultimele două au aceeași arie dacă conturul este pe deplin simetric. Însă dacă nu este îndeplinită condiția de simetrie, atunci numai una dă minimul absolut al ariei, în timp ce cealaltă va da un minim relativ, cu condiția ca minimul să fie căutat printre suprafețele simplu conexe. Posibilitatea soluției de genul

1 depinde de faptul că, admițînd suprafețe de genul 1, putem obține o arie mai mică decît dacă impunem condiția ca suprafața să fie simplu conexă. Deformînd conturul, dacă deformarea este destul de puternică, trebuie să ajungem la un punct în care acest lucru nu mai este adevărat. În acel moment, suprafața de genul 1 devine din ce în ce mai instabilă și sare brusc, în mod discontinuu, în soluția stabilă simplu conexă, reprezentată în fig. 245 sau fig. 246. Dacă începem cu una dintre aceste soluții simplu conexe, ca de pildă fig. 246, o putem deforma astfel, încît cealaltă soluție simplu conexă din fig. 245 să devină mult mai stabilă. Consecința este că, la un anumit moment, va avea loc o trecere discontinuă de la una la alta. Inversînd încet deformarea, revenim la poziția inițială a conturului, dar cu cealaltă soluție. Putem repeta procesul în sens opus și putem oscila în acest mod prin treceri discontinue între cele două tipuri. Procedînd cu grijă, putem transforma, de asemenea, în mod discontinuu, oricare dintre soluțiile simplu conexe, în cea de genul 1. În acest scop trebuie să aducem părțile în formă de disc foarte aproape una de alta, astfel încît suprafața de genul 1 să devină mult mai stabilă. Uneori în acest proces apar mai întîi bucăți intermediare de peliculă, care trebuie distruse, pentru a obține suprafața de genul 1.

Acest exemplu demonstrează nu numai posibilitatea diferitelor soluții de același tip topologic, dar și a unor tipuri diferite în același contur; mai mult, el ilustrează din nou posibilitatea unor treceri discontinue de la o soluție la alta, în timp ce condițiile problemei se modifică în mod continuu. Este ușor de construit modele mai complicate de aceeași natură și de studiat experimental comportarea lor.

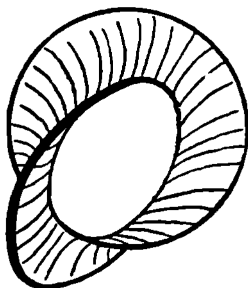


Fig. 248. Curbe înlănțuite

Un fenomen interesant este apariția suprafețelor minima, mărginite de două sau mai multe curbe închise înlănțuite. Pentru cazul a două cercuri, obținem suprafața arătată în fig. 248. Dacă în acest exemplu, cercurile sînt perpendiculare și dacă dreapta de intersecție a planelor lor este un diametru în ambele cercuri, atunci vor exista două forme simetric opuse ale acestei supra-

fețe, eu aceeași arie. Dacă deplasăm acum cercurile puțin unul față de altul forma va varia în mod continuu, cu toate că pentru fiecare poziție numai una este un minim absolut, iar cealaltă este un minim relativ. Dacă cercurile sînt mișcate astfel, încît să se formeze minimul relativ, el va sări la un anumit moment în poziția de minim absolut. Aici ambele suprafețe minima posibile au același caracter topologic, ca și suprafețele din fig. 245—246, dintre care una poate fi făcută să sară în cealaltă printr-o deformare ușoară a conturului.

4. Soluții experimentale ale altor probleme matematice

Datorită acțiunii tensiunii superficiale, o peliculă de lichid se află în echilibru stabil numai dacă aria ei este minimă. Aceasta este o sursă inepuizabilă de experiențe semnificative din punct de vedere matematic. Dacă unele părți ale frontierei unei pelicule se pot deplasa liber pe suprafețe date, ca de pildă pe anumite plane, atunci pe aceste frontiere, pelicula va fi perpendiculară pe suprafața dată.

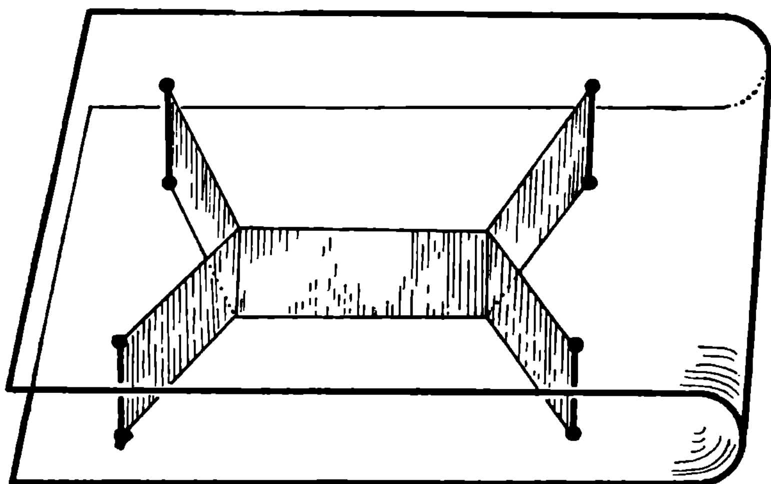


Fig. 249. Demonstrarea celei mai scurte legături între patru puncte

Putem folosi acest fapt pentru a da demonstrații surprinzătoare problemei lui Steiner și generalizărilor ei (cf. § 5). Două plăci paralele, confecționate din stielă sau din material plastic transparent, sînt unite prin trei sau mai multe bare perpendiculare. Dacă scufundăm acest obiect într-o soluție de săpun, scoțîndu-l apoi, pelicula formează un sistem de plane verticale, cuprinse

între cele două plăci şi care unesc barele fixe. Proiecţia care apare pe plăcile de sticlă constituie soluţia problemei discutată la pp. 395—396.

Dacă plăcile nu sînt paralele, atunci barele nu sînt perpendiculare pe ele, sau dacă plăcile sînt curbate, atunci curbele formate de peliculă pe plăci nu vor fi rectilinii şi vor ilustra noi probleme variaţionale.

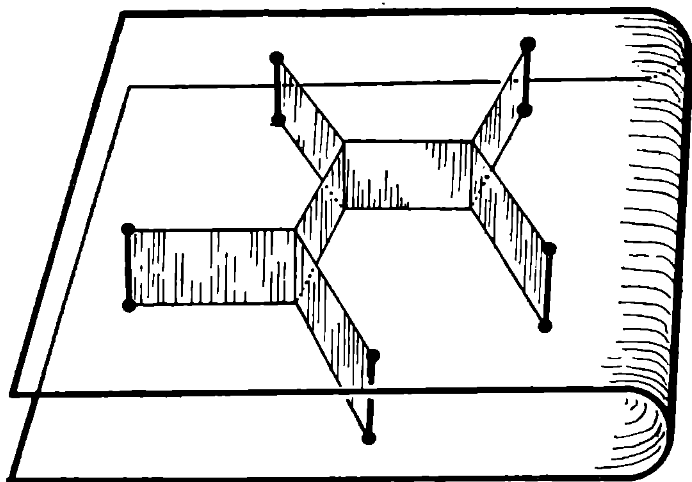


Fig. 250. Cea mai scurtă legătură între cinci puncte

Apariţia curbelor în lungul cărora se întîlnesc trei foi ale unei suprafeţe minima, care formează unghiuri de 120°, poate fi privită ca generalizare la mai multe dimensiuni, a fenomenului legat de problema lui Steiner. Acest lucru devine limpede, de pildă dacă unim două puncte A, B din spaţiu prin trei curbe şi dacă studiem sistemul stabil corespunzător de pelicule de săpun. Drept cel mai simplu caz îl vom lua pe acela în care una dintre curbe este segmentul de dreaptă AB, iar celelalte două sînt arce de cerc egale. Rezultatul este arătat în fig. 251. Dacă planele arcelor formează un unghi mai mic de 120°, obţinem trei suprafeţe care se întîlnesc formînd unghiuri de 120°; dacă întoarcem cele două arce, mărind unghiul dintre ele, soluţia variază continuu, transformîndu-se în două segmente circulare plane.

Să unim acum pe A şi B prin trei curbe mai complicate. Ca exemplu, putem lua trei linii frînte, fiecare fiind formată din trei muchii ale aceluiaşi cub şi care unesc două vîrfuri diagonal opuse: obţinem trei suprafeţe egale, care se întîlnesc în diagonala cubului. (Obţinem acest sistem de suprafeţe din acelea arătate în fig. 240, distrugînd peliculele adiacente la trei muchii alese în mod potrivit.) Dacă facem ca cele trei linii frînte, care unesc pe A cu B să fie

mobile, putem vedea că linia de intersecție triplă se curbează. Unghiurile vor fi tot de 120° (fig. 252).

Toate fenomenele în care se întîlnesc trei suprafețe minima după anumite curbe sînt esențial de natură asemănătoare. Ele sînt generalizări ale problemei plane, de unire a n puncte, prin cel mai scurt sistem de drumuri.

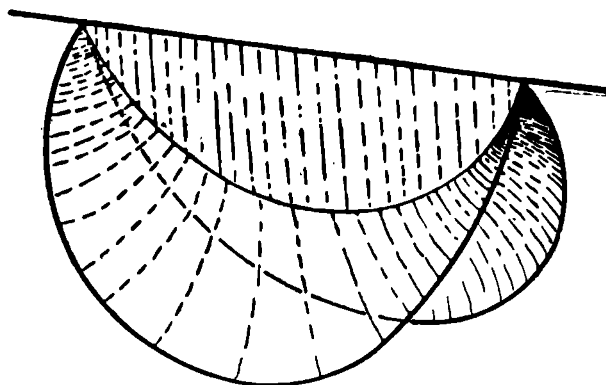


Fig. 251. Trei suprafețe care se întîlnesc sub unghiuri de 120°, subîntinse de trei sîrme care unesc două puncte

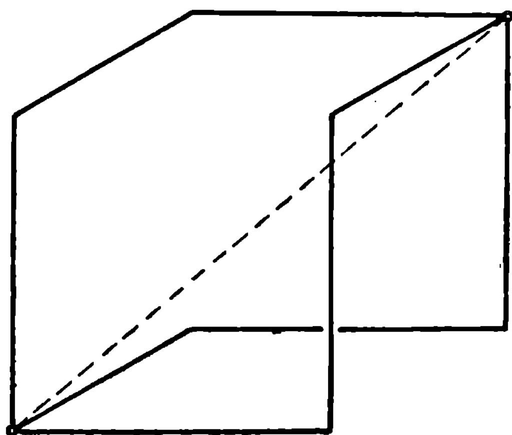


Fig. 252. Trei linii frînte care unesc două puncte

În sfîrşit, cîteva cuvinte referitoare la baloanele de săpun. Baloanele de săpun sferice arată că dintre toate suprafețele închise, care includ un anumit volum (determinat de cantitatea de aer conținută în interior), sfera are cea mai mică arie. Dacă considerăm baloane de săpun de volum dat, care tind

să se contracteze spre o arie minimă, dar care sînt restrînse prin anumite condiții, atunci suprafețele care rezultă nu vor fi sfere, ci suprafețe de curbură medie constantă, dintre care sferele și cilindrii circulari sînt exemple particulare.

De exemplu, să suflăm un balon de săpun între două plăci paralele de sticlă, care au fost umezite în prealabil cu soluția de săpun. Atunci cînd balonul

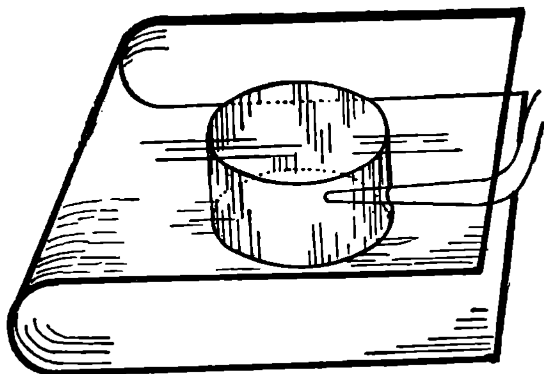


Fig. 253. Demonstrarea proprietății izoperimetrice a cercului

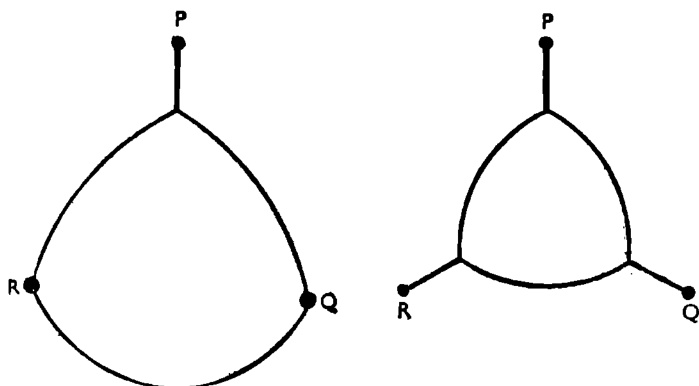


Fig. 254—255. Figuri izoperimetrice cu restricții la frontieră

atinge una dintre plăci, el ia brusc forma unei emisfere; îndată ce el atinge cealaltă placă, el ia forma unui cilindru circular, demonstrînd astfel în modul cel mai frapant proprietatea izoperimetrică a cercului. Faptul că pelicula de săpun aderă vertical la suprafața care o mărginește, este cheia acestei expe-

rienţe. Suflînd baloane de săpun între două plăci, cu bare perpendiculare de legătură, putem ilustra problemele discutate la pp. 395—396.

Putem studia comportarea soluţiei problemei izoperimetrice mărind sau micşorînd cantitatea de aer din balon, cu ajutorul unui tub cu vîrf ascuţit. Aspirînd însă aerul, nu obţinem figurile de la p. 396, care constau din arce circulare tangente două cîte două. Pe măsură ce volumul de aer cuprins descreşte, unghiurile triunghiului circular nu vor descreşte (teoretic) sub 120°; obţinem formele arătate în fig. 254—255, care tind din nou spre segmente de dreaptă ca în fig. 235, pe măsură ce aria tinde spre zero. Motivul matematic pentru care peliculele de săpun nu formează arce tangente este faptul că îndată ce balonul se separă de vîrfuri, curbele de legătură nu mai trebuie să fie socotite de două ori. Experienţele corespunzătoare sînt ilustrate de fig. 256 şi fig. 257.

Exerciţiu: Studiaţi problema matematică care să corespundă următoarelor condiţii: să se găsească un triunghi circular care cuprinde o arie dată, astfel încît perimetrul său plus cele trei segmente care unesc vîrfurile cu punctele date să aibă o lungime minimă.

Un contur cubic, în interiorul căruia umflăm un balon de săpun, va da suprafeţe de curbură medie constantă, ale căror baze sînt pătrate, dacă balonul depăşeşte conturul. Pe măsură ce extragem aer din balon, aspirîndu-l printr-un

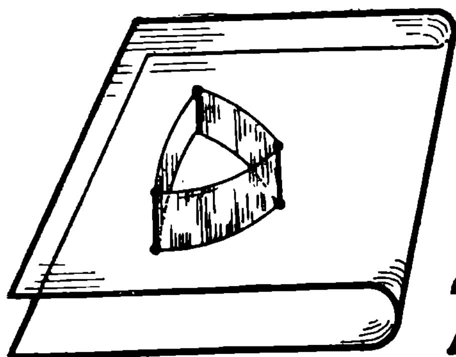


Fig. 256.

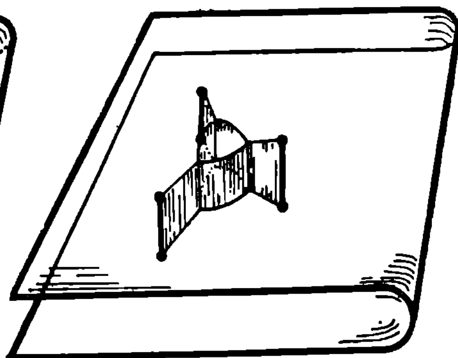


Fig. 257.

pai, obţinem un şir de figuri frumoase, care se reduc la cea din fig. 258. Fenomenele de stabilitate şi de tranziţie între diferitele stări de echilibru constituie o sursă de experienţe foarte instructive din punct de vedere matematic. Experienţele ilustrează teoria valorilor staţionare, deoarece stările de tranziţie pot fi făcute să aibă loc în aşa fel, încît să treacă printr-un echilibru instabil, care este o „stare staţionară".

De exemplu, structura cubică din fig. 240 manifestă o asimetrie, în cazul în care un plan vertical din centru uneşte cele 12 suprafeţe care pornesc din

muchii. Deci trebuie să existe cel puțin alte două poziții de echilibru, una cu un pătrat central vertical și una cu un pătrat central orizontal. De fapt, suflînd aer printr-un tub subțire în muchiile acestui pătrat, putem face ca pătratul să se reducă la un punct, centrul cubului; această poziție de echilibru instabil va trece imediat într-una dintre celelalte poziții stabile, obținută din cea inițială printr-o rotație de 90°.

O experiență asemănătoare poate fi făcută cu o peliculă de săpun care ilustrează problema lui Steiner, în cazul a patru puncte care formează un pătrat (fig. 219—220).

Dacă dorim să obținem soluțiile acestor probleme drept cazuri limită ale unor probleme izoperimetrice — de exemplu, dacă dorim să obținem fig. 240 din fig. 258 —, trebuie să aspirăm aerul din balon. Acum fig. 258 este complet simetrică și limita ei, atunci cînd conținutul balonului dispare, va fi un sistem simetric, format din 12 plane care se întîlnesc în centru. Acest

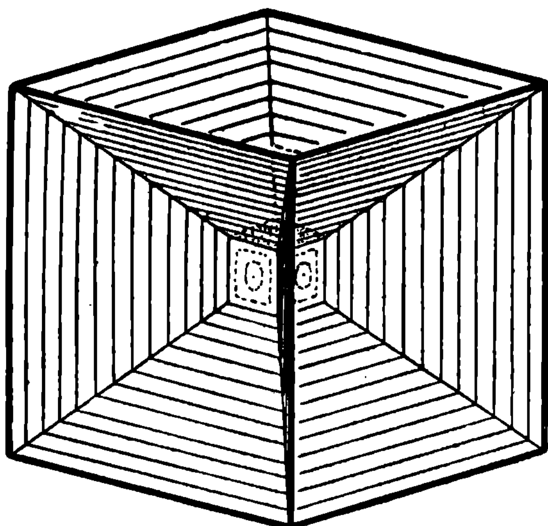


Fig. 258.

lucru poate fi observat în realitate, însă poziția obținută la limită nu este un echilibru stabil; ea va trece într-una dintre pozițiile din fig. 240. Folosind un lichid mai vîscos decît acela descris mai sus, întregul fenomen poate fi observat cu ușurință. El exemplifică faptul că chiar pentru probleme fizice, nu este necesar ca soluția unei probleme să depindă continuu de datele ei. Într-adevăr, în cazul limită al volumului nul, soluția dată în fig. 240 nu este limita soluției date de fig. 258 pentru volumul ε, atunci cînd ε tinde spre zero.

CALCULUL DIFERENŢIAL ŞI INTEGRAL

INTRODUCERE

Cu o ultrasimplificare absurdă, „inventarea” calculului diferenţial şi integral este atribuită uneori doar lui Newton şi Leibniz. În realitate, el este rezultatul unei îndelungate evoluţii, care nu a fost nici începută şi nici terminată de Newton şi Leibniz, dar în care amîndoi au jucat un rol decisiv. Un grup de matematicieni entuziaşti din diferite ţări din Europa secolului al XVII-lea se străduiau să continue opera matematică a lui Galilei şi Kepler. Prin corespondenţă şi călătorii, aceşti oameni menţineau un contact strîns. Două probleme centrale le atrăgeau atenţia. Prima era *problema tangentelor*: să se determine tangentele la o curbă dată, problema fundamentală a calculului diferenţial. A doua, *problema cvadraturii*: să se determine aria cuprinsă de o curbă dată, problema fundamentală a calculului integral. Marele merit al lui Newton şi Leibniz constă în înţelegerea clară a *legăturii intime dintre aceste două probleme*. În mîinile lor, aceste noi metode unificate au devenit instrumente puternice ale ştiinţei. O mare parte a succesului se datora minunatei notaţii simbolice inventate de Leibniz. Realizarea lui nu era nicidecum micşorată prin faptul că era legată de idei vagi şi de nesusţinut, capabile să perpetueze absenţa unei înţelegeri precise în minţile care preferă misticismul în locul clarităţii. Newton, cel mai mare savant dintre ei, pare să fi fost inspirat în cea mai mare măsură de Barrow (1630—1677), profesorul şi predecesorul său la Cambridge. Leibniz era mai degrabă un diletant. Strălucit avocat, diplomat şi filozof, una dintre cele mai active şi mobile minţi ale secolului său, Leibniz a învăţat noua matematică într-un timp extrem de scurt, de la fizicianul Huygens, în timp ce vizita Parisul, fiind în misiune diplomatică. Curînd după aceea, el a publicat rezultate care conţin nucleul calculului diferenţial şi integral modern. Newton, ale cărui descoperiri au fost făcute cu mult înainte, era un adversar al publicării. Mai mult, cu toate că el a găsit iniţial multe din rezultatele originale în capodopera sa *Principia* prin metodele calculului diferenţial şi integral, el a preferat să le prezinte în stilul geometriei clasice şi în *Principia* nu apare explicit aproape nici o urmă a calculului diferenţial şi integral. Lucrările sale referitoare la metoda

„fluxiunilor“ au fost publicate mult mai tîrziu. Curînd admiratorii săi au început o luptă îndîrjită cu prietenii lui Leibniz, pentru stabilirea priorității. Ei l-au acuzat pe Leibniz de plagiat, cu toate că, într-o atmosferă saturată de elementele unei noi teorii, nimic nu este mai natural decît descoperirea simultană și independentă. Cearta care a urmat pentru stabilirea priorității în ,,inventarea'' calculului diferențial și integral constituie un exemplu nefericit de exagerare a problemelor priorității și pretențiilor de proprietate intelectuală, capabile să otrăvească atmosfera contactului științific natural.

În analiza matematică a secolului al XVII-lea și a celei mai mari părți a secolului al XVIII-lea, idealul grec al raționamentului clar și riguros pare să fi fost abandonat. ,,Intuiția'' și ,,instinctul'' au înlocuit raționamentul în multe împrejurări importante. Acest fapt a constituit o încurajare a credinței necritice în puterea supraomenească a noilor metode. Se gîndea, în general, că o prezentare clară a rezultatelor calculului nu numai că nu era necesară, dar era chiar imposibilă. Dacă noua știință nu s-ar fi aflat în mîinile unui grup restrîns de oameni extrem de competenți, s-ar fi putut comite erori serioase. Acești pionieri erau ghidați de un puternic simț instinctiv, care îi apăra de erori. Dar atunci cînd revoluția franceză a deschis calea unei extinderi imense a învățămîntului superior, atunci cînd din ce în ce mai mulți oameni au vrut să participe la activitatea științifică, revizuirea critică a noii analize n-a mai putut fi amînată. Această problemă a fost rezolvată în secolul al XIX-lea și astăzi analiza poate fi predată fără nici o urmă de mister și cu deplină rigoare. Nu mai este nici un motiv ca acest instrument fundamental al științelor să nu fie înțeles de orice persoană instruită.

Scopul acestui capitol este de a da o introducere elementară, în care accentul este pus pe înțelegerea noțiunilor fundamentale și mai puțin pe aspectul formal al regulilor de lucru. Limbajul intuitiv va fi folosit peste tot, dar întotdeauna într-un mod compatibil cu noțiunile precise și cu procedeul clar.

§ 1. INTEGRALA

1. Aria ca limită

Pentru a calcula aria unei figuri plane, alegem ca *unitate de arie* un pătrat, ale cărui laturi sînt egale cu unitatea de lungime. Dacă unitatea de lungime este centimentrul, unitatea corespunzătoare de arie va fi centimetrul pătrat, adică pătratul ale cărui laturi sînt egale cu un centimetru. Pe baza acestei definiții, se poate calcula foarte ușor aria unui dreptunghi. Dacă p și q sînt lungimile a două laturi adiacente, măsurate în raport cu unitatea de lungime, atunci aria dreptunghiului este de pq unități pătrate sau, mai scurt, aria este egală cu produsul pq. Acest lucru este adevărat pentru p și q arbitrare, raționale sau nu. Dacă p și q sînt raționale, obținem acest rezultat scriind $p = m/n$, $q = m'/n'$, unde m, n, m', n' sînt întregi. Apoi găsim măsura comună $1/N = 1/nn'$ a celor două laturi, astfel încît $p = mn' \cdot 1/N$, $q = nm' \cdot 1/N$. În

sfîrşit, subdivizăm dreptunghiul în pătrate mici de latură $1/N$ şi de arie $1/N^2$. Numărul acestor pătrate este egal cu $nm' \cdot mn'$, iar aria totală este egală cu $nm'mn' \cdot 1/N^2 = nm'mn'/n^2n'^2 = m/n \cdot m'/n' = pq$. Dacă p şi q sînt iraţionale, obţinem acelaşi rezultat, înlocuind mai întîi pe p şi q, respectiv cu numerele raţionale p_r şi q_r care le aproximează, şi apoi făcînd ca p_r şi q_r să tindă spre p şi q.

Este evident, din punct de vedere geometric, că aria unui triunghi este egală cu jumătate din aria unui dreptunghi, care are aceeaşi bază b şi înălţimea h; deci aria unui triunghi este dată de expresia cunoscută $(1/2)bh$. Orice domeniu din plan, mărginit de una sau mai multe linii poligonale, poate fi descompus în triunghiuri; de aceea, aria lui poate fi obţinută ca sumă a ariilor acestor triunghiuri.

Necesitatea unei metode mai generale pentru calculul ariilor apare atunci cînd căutăm aria unei figuri care nu mai este mărginită de poligoane, ci de *curbe*. Cum trebuie să determinăm, de pildă, aria unui disc circular sau a unui segment de parabolă? Această problemă fundamentală, care se află la baza calculului integral, a fost tratată încă din secolul al III-lea î.e.n. de către Arhimede, care a calculat astfel de arii prin procedeul „exhaustiunii". Împreună cu Arhimede şi cu marii matematicieni care au trăit pînă în vremea lui Gauss, putem adopta atitudinea „naivă", în baza căreia ariile curbilinii sînt entităţi date intuitiv şi că nu se pune problema *definirii*, ci numai a *calculării* lor (vedeţi însă discuţia de la p. 483). În domeniul nostru înscriem un domeniu de aproximare cu frontieră poligonală şi cu o arie bine definită. Alegînd un alt domeniu poligonal care include pe primul, obţinem o aproximare mai bună a domeniului dat. Procedînd în acest mod putem „epuiza" treptat întreaga arie şi obţinem aria domeniului dat, ca limită a ariilor unui şir de domenii poligonale înscrise, cu un număr crescător de laturi alese în mod potrivit. Aria unui cerc de rază 1 poate fi calculată în acest fel, iar valoarea ei numerică se notează cu ajutorul simbolului π.

Arhimede a aplicat acest procedeu general pentru cerc şi pentru segmentul de parabolă. În decursul secolului al XVII-lea au fost tratate cu succes multe alte cazuri. De fiecare dată calculul efectiv al limitei depindea de un artificiu ingenios, adaptat în mod special problemei considerate. Una dintre marile realizări ale calculului diferenţial şi integral a fost înlocuirea acestor procedee speciale şi restrictive pentru calculul ariilor printr-o metodă generală şi puternică.

2. Integrala

Prima noţiune fundamentală a calculului[1] este aceea de integrală. În această secţiune vom înţelege integrala ca expresie a *ariei aflate sub o curbă*, arie definită cu ajutorul unei limite. Dacă este dată o funcţie continuă pozitivă

[1] diferenţial şi integral (v. nota de la p. 454)

$y = f(x)$, de pildă $y = x^2$ sau $y = 1 + \cos x$, să considerăm domeniul mărginit inferior de un segment aflat pe axa Ox şi cuprins între o coordonată a şi o coordonată mai mare b, de perpendicularele ridicate în aceste puncte pe axa Ox, şi mărginit superior de curba $y = f(x)$. Scopul nostru este de a calcula aria A a acestui domeniu.

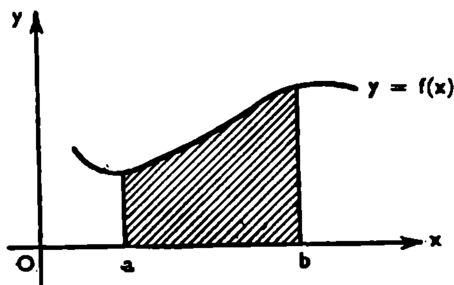


Fig. 259. Integrala ca arie

Deoarece un astfel de domeniu, în general, nu poate fi descompus în dreptunghiuri sau triunghiuri, nu dispunem de o expresie imediată a ariei A, prin calcul explicit. Putem însă găsi o valoare aproximativă a lui A şi, prin urmare, putem reprezenta pe A ca limită în modul următor: subdivizăm intervalul cuprins între $x = a$ şi $x = b$ într-un număr de subintervale mici, ridicăm perpendiculare în fiecare punct al subdiviziunii şi înlocuim fiecare fîşie a domeniului aflat sub curbă printr-un dreptunghi, a cărui înălţime este aleasă undeva între cea mai mare şi cea mai mică înălţime a curbei în acea fîşie. Suma S a ariilor acestor dreptunghiuri dă o valoare aproximativă a ariei reale A, aflată sub curbă. Precizia acestei aproximări va fi cu atît mai bună cu cît numărul dreptunghiurilor este mai mare şi cu cît lăţimea fiecărui dreptunghi este mai mică. Astfel, putem caracteriza aria care ne interesează printr-o limită: dacă formăm un şir

$$(1) \qquad S_1,\ S_2,\ S_3,\ \ldots,$$

de aproximări dreptunghiulare ale ariei aflate sub curbă, astfel încît lăţimea celui mai lat dreptunghi din S_n să tindă către 0, cînd n creşte, atunci şirul (1) tinde spre limita A:

$$(2) \qquad S_n \to A,$$

iar această limită A, adică aria aflată sub curbă, este independentă de modul particular în care este ales şirul (1), cu condiţia ca lăţimile dreptunghiurilor de aproximare să tindă spre zero. (De exemplu, S_n poate proveni din S_{n-1} prin adăugarea unuia sau mai multor puncte de subdiviziune acelora

care definesc pe S_{n-1}, sau alegerea punctelor de subdiviziune pentru S_n poate fi cu totul independentă de alegerea făcută pentru S_{n-1}.) Aria A a domeniului, exprimată prin acest proces de trecere la limită, este prin definiţie *integrala funcţiei* f(x) *de la* a *la* b. Introducînd simbolul special, ,,semnul integral'', ea se scrie astfel

$$A = \int_a^b f(x)\,dx. \tag{3}$$

Simbolul $\int$, ,,dx'' şi numele ,,integrală'' au fost introduse de Leibniz, pentru a sugera modul în care se obţine limita. Pentru a explica această notaţie vom repeta mai detaliat procesul aproximării ariei A. În acelaşi timp, formularea analitică a procesului de trecere la limită va face posibilă renunţarea la ipotezele respective $f(x) \geqslant 0$ şi $b > a$ şi, în sfîrşit, vom putea elimina noţiunea iniţială intuitivă de arie, ca bază a definiţiei integralei (aceasta se va face în supliment, § 1).

Să subdivizăm intervalul cuprins între a şi b, în n subintervale mici, care, pentru simplificare, vor fi presupuse de lungime egală $(b - a)/n$. Să notăm punctele de subdiviziune prin

$$x_0 = a, \quad x_1 = a + \frac{b-a}{n},$$

$$x_2 = a + \frac{2(b-a)}{n}, \; \ldots, \; x_n = a + \frac{n(b-a)}{n} = b.$$

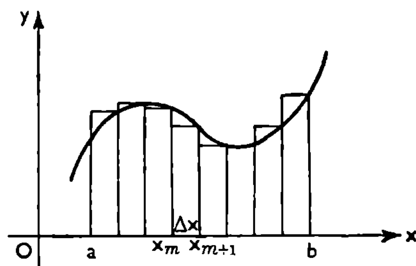


Fig. 260. Aproximarea ariei prin dreptunghiuri înguste

Pentru cantitatea $(b - a)/n$, diferenţa dintre valorile consecutive ale lui x, introducem notaţia Δx (se citeşte ,,delta x''),

$$\Delta x = \frac{b-a}{n} = x_{j+1} - x_j,$$

unde simbolul Δ înseamnă, pur şi simplu, ,,diferenţa'' (este un simbol ,,operator'' şi nu trebuie confundat cu un număr). Putem alege înălţimea fiecărui

drpetunghi de aproximare egală cu valoarea $y = f(x)$ în extremitatea din dreapta a subintervalului. Atunci suma ariilor acestor dreptunghiuri va fi:

$$S_n = f(x_1) \cdot \Delta x + f(x_2) \cdot \Delta x + \ldots + f(x_n) \cdot \Delta x \tag{4}$$

sau prescurtat

$$S_n = \sum_{j=1}^{n} f(x_j) \cdot \Delta x. \tag{5}$$

Simbolul $\sum_{j=1}^{}$ (se citeşte „sigma de la $j = 1$ la n") reprezintă suma tuturor expresiilor care se obţin cînd j ia pe rînd valorile 1, 2, 3, ..., n.

Folosirea simbolului Σ pentru a exprima sub o formă concisă rezultatul unei adunări poate fi ilustrată prin exemplele următoare:

$$2 + 3 + 4 + \ldots + 10 = \sum_{j=2}^{10} j,$$

$$1 + 2 + 3 + \ldots + n = \sum_{j=1}^{n} j,$$

$$1^2 + 2^2 + 3^2 + \ldots + n^2 = \sum_{j=1}^{n} j^2,$$

$$aq + aq^2 + \ldots + aq^n = \sum_{j=1}^{n} aq^j,$$

$$a + (a + d) + (a + 2d) + \ldots + (a + nd) = \sum_{j=0}^{n} (a + jd).$$

Acum să formăm un şir de astfel de aproximări S_n, în care n creşte indefinit, astfel încît numărul termenilor din fiecare sumă (5) să crească, în timp ce fiecare termen al unei sume, de forma $f(x_j)\Delta x$, tinde spre 0, datorită factorului $\Delta x = (b - a)/n$. Cînd n creşte, această sumă tinde spre aria A:

$$A = \lim \sum_{j=1}^{n} f(x_j)\ \Delta x = \int_a^b f(x)\, dx. \tag{6}$$

Leibniz a simbolizat această trecere la limită de la sumele de aproximare S_n la A, înlocuind semnul de sumare Σ prin $\int$, iar simbolul diferenţei Δ, prin d. (Simbolul de sumare Σ se scria de obicei S, pe vremea lui Leibniz, iar simbolul $\int$ reprezintă un S stilizat.) Cu toate că simbolismul lui Leibniz

sugerează foarte bine modul în care se obţine integrala, ca limită a unei sume finite, trebuie să avem grijă să nu acordăm o însemnătate prea mare unei pure convenţii şi modului în care trebuie să fie notată limita. În primele zile ale analizei, cînd noţiunea de limită nu era clar înţeleasă şi nu se afla întotdeauna prezentă în minte, se explica înţelesul integralei, spunîndu-se că „diferenţa finită Δx este înlocuită prin cantitatea infinit de mică dx, iar integrala este suma unei infinităţi de cantităţi infinit mici $f(x)dx$". Cu toate că „infinitul mic" are o anumită atracţie pentru minţile înclinate spre speculaţii filozofice, el nu-şi mai găseşte locul în matematica modernă. Nu servim nici unui scop util, înconjurînd noţiunea clară de integrală cu un nor de fraze fără sens. Leibniz însuşi era derutat uneori de puterea sugestivă a simbolurilor sale; într-adevăr, ele operează ca şi cum ar desemna o sumă de cantităţi „infinit mici" cu care totuşi se poate opera, într-o oarecare măsură, ca şi cu cantităţile obişnuite. De fapt, cuvîntul „integral" a fost dat pentru a indica faptul că întreaga arie, sau aria integrală A, este formată din părţile „infinitesimale" $f(x)dx$. În orice caz, au trebuit să treacă aproape o sută de ani după Newton şi Leibniz, înainte ca să se înţeleagă limpede că noţiunea de limită este singura bază adevărată pentru definirea integralei. Stînd ferm pe această bază, putem evita toate confuziile, toate dificultăţile şi tot nonsensul atît de tulburător de la începutul dezvoltării analizei.

3. Observaţii generale asupra noţiunii de integrală. Definiţia generală

În definiţia geometrică a integralei ca arie, am presupus în mod explicit că $f(x)$ nu este nicăieri negativă pe intervalul de integrare $[a, b]$, adică nici o porţiune a graficului nu se află sub axa Ox. Dar, în definiţia analitică a integralei, ca limită a unui şir de sume S_n, această ipoteză este inutilă. Luăm

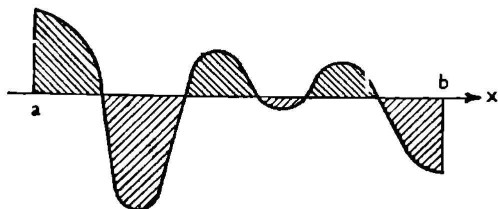


Fig. 261. Arii pozitive şi arii negative

pur şi simplu micile cantităţi $f(x_j) \cdot \Delta x$, formăm suma lor şi trecem la limită; acest procedeu îşi păstrează semnificaţia chiar şi dacă unele sau toate valorile $f(x_j)$ sînt negative. Interpretînd geometric acest lucru cu ajutorul ariilor (fig. 261), găsim că integrala lui $f(x)$ este suma *algebrică* a ariilor mărginite de

grafic şi de axa Ox, unde ariile aflate sub axa Ox sînt considerate negative, iar celelalte pozitive.

Se poate întîmpla ca în aplicaţii să fim conduşi la integrale $\int_a^b f(x)\mathrm{d}x$, unde b este mai mic decît a, astfel încît $(b - a)/n = \Delta x$ este un număr negativ. În definiţia analitică avem deci $f(x_j) \cdot \Delta x$ negativ, dacă $f(x_j)$ este pozitiv şi Δx este negativ etc. Cu alte cuvinte, valoarea integralei va fi opusul valorii integralei de la b la a. Obţinem astfel următoarea regulă simplă a integralei:

$$\int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\mathrm{d}x.$$

Mai departe trebuie să subliniem faptul că valoarea integralei rămîne neschimbată, chiar dacă nu ne limităm la puncte echidistante x_j de subdiviziune sau, ceea ce este acelaşi lucru, la diferenţe $\Delta x = x_{j+1} - x_j$ egale. Putem alege pe x_j în alte moduri, astfel încît diferenţele $\Delta x_j = x_{j+1} - x_j$ să nu fie egale (deci, trebuie să le deosebim prin indici). Chiar şi în această ipoteză, sumele

$$S_n = f(x_1)\,\Delta x_0 + f(x_2)\Delta x_1 + \ldots + f(x_n)\,\Delta x_{n-1}$$

precum şi sumele

$$S'_n = f(x_0)\,\Delta x_0 + f(x_1)\Delta x_1 + \ldots + f(x_{n-1})\Delta x_{n-1}$$

vor tinde spre aceeaşi limită, valoarea integralei $\int_a^b f(x)\mathrm{d}x$, dacă avem grijă ca atunci cînd n creşte, toate diferenţele $\Delta x_j = x_{j+1} - x_j$ să tindă spre zero, astfel încît cea mai mare diferenţă de acest fel, pentru un n dat, să tindă spre zero, cînd n creşte.

În consecinţă, *definiţia finală a integralei* este dată de

$$\int_a^b f(x)\mathrm{d}x = \lim \sum_{j=1}^{n} f(v_j)\,\Delta x_j, \tag{6a}$$

cînd $n \to \infty$. În această limită, v_j poate reprezenta orice punct din intervalul $x_j \leqslant v_j \leqslant x_{j+1}$ şi singura restricţie impusă subdiviziunii este ca cel mai lung interval, de lungime $\Delta x_j = x_{j+1} - x_j$, să tindă spre zero, cînd n creşte.

Existența limitei (6a) nu trebuie demonstrată, dacă admitem, de la început, noțiunea de arie aflată sub o curbă și posibilitatea aproximării acestei arii prin sume de dreptunghiuri. Totuși, așa cum se va vedea mai departe (p. 483), o analiză mai atentă arată că este de dorit și chiar necesar, pentru o prezen-

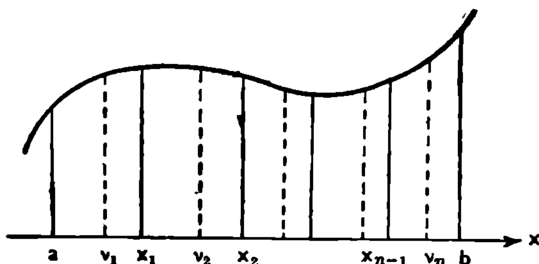


Fig. 262. O subdiviziune arbitrară în definiția generală a integralei

tare logică completă a noțiunii de integrală, să se demonstreze existența limitei pentru orice funcție continuă $f(x)$, fără a recurge la noțiunea geometrică inițială de arie.

4. Exemple de integrare. Integrarea lui x^r

Pînă acum, discuția noastră asupra integralei a fost pur teoretică. Problema fundamentală este dacă schema generală de formare a sumelor S_n, urmată de trecerea la limită, duce într-adevăr la rezultate palpabile în cazurile concrete. Desigur, aceasta va necesita un raționament suplimentar, legat de funcția particulară $f(x)$, a cărei integrală trebuie găsită. Atunci cînd Arhimede, cu 2 000 de ani în urmă, a găsit aria segmentului de parabolă, el a efectuat ceea ce astăzi se numește integrarea funcției $f(x) = x^2$, printr-un artificiu foarte ingenios; în secolul al XVII-lea, precursorii analizei moderne au reușit să rezolve probleme de integrare pentru funcții simple, cum ar fi x^n, tot prin artificii speciale. Numai după examinarea unui număr mare de cazuri particulare s-a găsit o metodă generală de atacare a problemei integrării, prin metodele sistematice ale analizei și astfel, cîmpul de rezolvare a problemelor particulare a fost considerabil lărgit. În capitolul de față vom discuta cîteva dintre problemele speciale instructive, care aparțin epocii „preanalizei", deoarece nimic nu poate ilustra mai bine integrarea ca proces de trecere la limită.

a) Începem cu un exemplu foarte banal. Dacă $y = f(x)$ este o constantă, de exemplu $f(x) = 2$, atunci, evident, integrala $\int_a^b 2\mathrm{d}x$, înțeleasă ca arie, este

egală cu $2(b - a)$, deoarece aria unui dreptunghi este egală cu produsul dintre bază şi înălţime. Vom compara acest rezultat cu definiţia integralei (6), ca limită. Dacă substituim în (5) $f(x_j) = 2$, pentru toate valorile lui j găsim că

$$S_n = \sum_{j=1}^{n} f(x_j)\Delta x = \sum_{j=1}^{n} 2\Delta x = 2\sum_{j=1}^{n} \Delta x = 2(b - a)$$

pentru orice n, deoarece

$$\sum_{j=1}^{n} \Delta x = (x_1 - x_0) + (x_2 - x_1) + \ldots + (x_n - x_{n-1}) = x_n - x_0 = b - a.$$

b) Aproape tot atît de simplă este integrarea lui $f(x) = x$. În acest caz, $\int_a^b x dx$ este aria unui trapez (fig. 263) şi aceasta, după cum ştim din geometria elementară, este egală cu

$$(b - a)\frac{b + a}{2} = \frac{b^2 - a^2}{2}.$$

Acelaşi rezultat se obţine şi din definiţia (6) a integralei, după cum se vede prin trecerea la limită efectivă, fără a folosi figura geometrică. Dacă substituim $f(x) = x$ în (5), atunci suma S_n devine

$$S_n = \sum_{j=1}^{n} x_j \Delta x = \sum_{j=1}^{n} (a + j\Delta x)\Delta x =$$

$$= (na + \Delta x + 2\Delta x + 3\Delta x + \ldots + n\Delta x)\Delta x =$$

$$= na\Delta x + (\Delta x)^2(1 + 2 + 3 + \ldots + n).$$

Folosind formula (1) de la p. 28, care dă suma progresiei aritmetice $1 + + 2 + 3 + \ldots + n$, găsim

$$S_n = na\Delta x + \frac{n(n + 1)}{2}(\Delta x)^2.$$

Deoarece $\Delta x = \dfrac{b - a}{n}$, aceasta este egală cu:

$$S_n = a(b - a) + \frac{1}{2}(b - a)^2 + \frac{1}{2n}(b - a)^2.$$

Dacă facem acum ca n să tindă spre infinit, ultimul termen tinde spre zero și obținem

$$\lim S_n = \int_a^b x \, dx = a(b - a) + \frac{1}{2}(b - a)^2 = \frac{1}{2}(b^2 - a^2),$$

în conformitate cu interpretarea geometrică a integralei ca arie.

c) Mai puțin banală este integrarea funcției $f(x) = x^2$. Arhimede a folosit metode geometrice pentru a rezolva problema echivalentă a găsirii ariei unui segment al parabolei $y = x^2$. Vom proceda analitic, pe baza definiției (6a). Pentru simplificare alegem pe 0 ca „limită inferioară" a a integralei; atunci $\Delta x = b/n$. Deoarece $x_j = j \cdot \Delta x$ și $f(x_j) = j^2 (\Delta x)^2$, obținem pentru S_n expresia

$$S_n = \sum_{j=1}^{n} f(j\Delta x)\Delta x = [1^2 \cdot (\Delta x)^2 + 2^2 \cdot (\Delta x)^2 + \ldots + n^2(\Delta x)^2] \cdot \Delta x =$$

$$= (1^2 + 2^2 + \ldots + n^2)(\Delta x)^3.$$

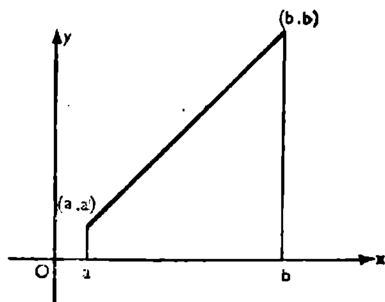


Fig. 263. Aria unui trapez

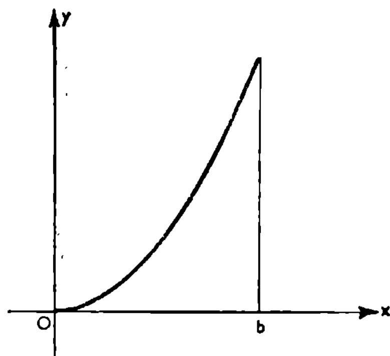


Fig. 264. Aria mărginită de o parabolă

Acum putem calcula limita. Folosind formula

$$1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

stabilită la p. 29 și făcînd substituția $\Delta x = b/n$, obținem

$$S_n = \frac{n(n+1)(2n+1)}{6} \cdot \frac{b^3}{n^3} = \frac{b^3}{6}\left(1 + \frac{1}{n}\right)\left(2 + \frac{1}{n}\right).$$

Această transformare preliminară face ca trecerea la limită să devină foarte simplă, deoarece $1/n$ tinde spre zero cînd n crește indefinit. Prin urmare obținem ca limită $(b^3/6) \cdot 1 \cdot 2 = b^3/3$ și deci rezultatul final are forma

$$\int_0^b x^2 \mathrm{d}x = b^3/3.$$

Aplicînd acest rezultat pentru aria cuprinsă între 0 și a avem

$$\int_0^a x^2 \mathrm{d}x = a^3/3,$$

și, prin scăderea ariilor,

$$\int_a^b x^2 \,\mathrm{d}x = (b^3 - a^3)/3.$$

Exercițiu: Demonstrați în același mod, folosind formula (5) de la p. 31 că

$$\int_a^b x^3 \mathrm{d}x = (b^4 - a^4)/4.$$

Stabilind formule generale pentru suma $1^k + 2^k + \ldots + n^k$ a puterilor k ale întregilor de la 1 la n, obținem rezultatul

$$\int_a^b x^k \mathrm{d}x = \frac{b^{k+1} - a^{k+1}}{k+1}, \tag{7}$$

k fiind un întreg pozitiv arbitrar.

* În loc de a proceda în acest mod, putem obține, pe o cale mai simplă, un rezultat mai general, folosind observația precedentă, în virtutea căreia putem calcula integrala cu ajutorul unor puncte de diviziune neechidistante. Vom stabili formula (7) nu numai pentru orice întreg pozitiv k, ci pentru un număr rațional pozitiv sau negativ arbitrar,

$$k = u/v,$$

unde u este un întreg pozitiv, iar v este un întreg pozitiv sau negativ. Doar valoarea $k = -1$, pentru care formula (7) nu are sens, este exclusă. Vom mai presupune că $0 < a < b$.

Pentru a obține formula integrală (7), să formăm sumele S_n, alegînd punctele de diviziune $x_0 = a, x_1, x_2, \ldots, x_n = b$, în *progresie geometrică*. Punem $\sqrt[n]{\frac{b}{a}} = q$, astfel încît $b/a =$

$= q^n$, şi să definim $x_0 = a$, $x_1 = aq$, $x_2 = aq^2$, ..., $x_n = aq^n = b$. Prin acest artificiu, după cum vom vedea, trecerea la limită devine foarte simplă. Pentru suma S_n, ţinînd seama că $f(x_j) = x_j^k = a^k q^{jk}$ şi $\Delta x_j = x_{j+1} - x_j = aq^{j+1} - aq^j$, găsim

$$S_n = a^k(aq - a) + a^k q^k(aq^2 - aq) + a^k q^{2k}(aq^3 - aq^2) + \\ + \dots + a^k q^{(n-1)k}(aq^n - aq^{n-1}).$$

Deoarece fiecare termen conţine factorii $a^k(aq - a)$, putem scrie

$$S_n = a^{k+1}(q - 1)\{1 + q^{k+1} + q^{2(k+1)} + \dots + q^{(n-1)(k+1)}\}.$$

Substituind pe q^{k+1} prin t, vedem că expresia cuprinsă în acolade este suma $1 + t + t^2 + + \dots + t^{n-1}$ a unei progresii geometrice care, în baza celor arătate la p.29, este egală cu $\frac{t^n - 1}{t - 1}$. Însă $t^n = q^{n(k+1)} = \left(\frac{b}{a}\right)^{k+1} = \frac{b^{k+1}}{a^{k+1}}$. Deci,

$$S_n = (q - 1)\left\{\frac{b^{k+1} - a^{k+1}}{q^{k+1} - 1}\right\} = \frac{b^{k+1} - a^{k+1}}{N}, \tag{8}$$

unde

$$N = \frac{q^{k+1} - 1}{q - 1}.$$

Pînă acum n a fost un număr fixat. Să-l facem acum pe n să crească şi să determinăm limita spre care tinde N. Cînd n creşte, rădăcina $\sqrt[n]{\frac{b}{a}} = q$ va tinde spre 1 (cf.p.341) şi, de aceea, atît numărătorul, cît şi numitorul lui N vor tinde spre zero, ceea ce ne obligă la o anumită precauţie. Să presupunem, mai întîi, că k este un întreg pozitiv; atunci împărţirea cu $q - 1$ poate fi efectuată şi obţinem (cf. p. 29) $N = q^k + q^{k-1} + \dots + q + 1$. Dacă n creşte, q tinde spre 1 şi deci q^2, q^3 ..., q^k vor tinde de asemenea spre 1, astfel încît N tinde către $k + 1$. Însă aceasta arată că S_n tinde spre $\frac{b^{k+1} - a^{k+1}}{k + 1}$, ceea ce trebuia arătat.

Exerciţiu: Arătaţi că pentru orice $k \neq -1$, k raţional, rămîne valabilă aceeaşi formulă de trecere la limită, $N \to k + 1$, şi de aceea şi rezultatul (7) rămîne valabil. Daţi mai întîi demonstraţia, folosind modelul de mai sus, pentru cazul în care k este un întreg negativ. Atunci, dacă $k = u/v$, scrieţi $q^{1/v} = s$ şi

$$N = \frac{s^{(k+1)v} - 1}{s^v - 1} = \frac{s^{u+v} - 1}{s^v - 1} = \frac{s^{u+v} - 1}{s - 1} \bigg/ \frac{s^v - 1}{s - 1}.$$

Dacă n creşte, atît s, cît şi q tind către 1 şi de aceea cele două cîturi din membrul drept tind, respectiv, către $u + v$ şi v, ceea ce ne dă din nou $\frac{u + v}{v} = k + 1$ pentru limita lui N.

În § 5 vom vedea cum se poate înlocui această discuţie lungă şi oarecum artificială, prin metodele mai simple şi mai puternice ale analizei.

Exerciţii: Verificaţi integrarea precedentă a lui x pentru cazurile $k = 1/2, -1/2, 2, -2, 3, -3$.

2) Găsiţi valorile integralelor:

$$\text{a)} \int_{-2}^{-1} x\,dx. \quad \text{b)} \int_{-1}^{+1} x\,dx. \quad \text{c)} \int_{1}^{2} x^2\,dx. \quad \text{d)} \int_{-1}^{-2} x^3\,dx. \quad \text{e)} \int_{0}^{n} x\,dx.$$

3) Găsiţi valorile integralelor:

$$\text{a)} \int_{-1}^{+1} x^3\,dx. \quad \text{b)} \int_{-2}^{2} x^3 \cos x\,dx. \quad \text{c)} \int_{-1}^{+1} x^4 \cos^2 x \sin^5 x\,dx. \quad \text{d)} \int_{-1}^{+1} \operatorname{tg} x\,dx.$$

(Indicaţie: Consideraţi graficele funcţiilor de sub semnul integral, ţineţi seama de simetria lor faţă de $x = 0$ şi interpretaţi integralele ca arii.)

*4) Integraţi pe $\sin x$ şi $\cos x$ de la 0 la b, făcînd substituţia $\Delta x = h$ şi folosind formulele de la p. 508.

5) Integraţi funcţiile $f(x) = x$ şi $f(x) = x^2$ de la 0 la b, divizînd intervalul în părţi egale şi folosind în formula (6a) valorile $v_j = (1/2)(x_j + x_{j+1})$.

*6) Folosind rezultatul (7) şi definiţia integralei cu valori egale pentru Δx, demonstraţi relaţia:

$$\frac{1^k + 2^k + \ldots + n^k}{n^{k+1}} \to \frac{1}{k+1}, \quad \text{cînd } n \to \infty.$$

$\left(\text{Indicaţie: puneţi } 1/n = \Delta x \text{ şi arătaţi că limita este egală cu } \int_0^1 x^k\,dx.\right)$

*7) Demonstraţi că pentru $n \to \infty$:

$$\frac{1}{\sqrt{n}}\left(\frac{1}{\sqrt{1+n}} + \frac{1}{\sqrt{2+n}} + \ldots + \frac{1}{\sqrt{n+n}}\right) \to 2(\sqrt{2} - 1).$$

(Indicaţie: Scrieţi această sumă astfel, încît limita ei să apară ca o integrală.)

8) Calculaţi aria unui segment parabolic, mărginit de un arc P_1P_2 şi de coarda P_1P_2 a parabolei $y = ax^2$, în funcţie de coordonatele x_1 şi x_2 ale celor două puncte.

5. Reguli pentru „calculul integral"

Un pas important în dezvoltarea analizei a fost făcut atunci cînd au fost formulate anumite reguli generale, cu ajutorul cărora problemele mai complicate au putut fi reduse la altele, mai simple, permiţînd astfel rezolvarea lor printr-un procedeu aproape mecanic. Caracterul algoritmic al acestor reguli este subliniat în special de notaţia lui Leibniz. Totuşi o concentrare prea mare asupra mecanismului rezolvării problemei poate degrada predarea analizei, reducînd-o la o simplă dexteritate.

Cîteva reguli simple de integrare rezultă imediat, fie din definiția (6), fie din interpretarea geometrică a integralelor ca arii.

Integrala sumei a două funcții este egală cu suma integralelor celor două funcții. Integrala produsului dintre o constantă c *și o funcție* f(x) *este egală cu produsul dintre constanta* c *și integrala funcției* f(x). Combinînd aceste două reguli se obține formula:

$$(9)\qquad \int_a^b [cf(x) + \mathrm{d}g(x)]\,\mathrm{d}x = c\int_a^b f(x)\,\mathrm{d}x + \mathrm{d}\int_a^b g(x)\,\mathrm{d}x.$$

Demonstrația rezultă imediat din definiția integralei ca limită a unei sume finite (5), deoarece formula corespunzătoare pentru o sumă S_n este evident adevărată. Regula se extinde imediat la suma de mai multe funcții.

Ca exemplu de folosire a acestei reguli, să considerăm un polinom

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n,$$

în care coeficienții $a_0, a_1, \dots, a_n$ sînt constante. Pentru a calcula integrala lui $f(x)$ de la a la b, integrăm, termen cu termen, conform regulei de mai sus. Folosind formula (7), găsim

$$\int_a^b f(x)\,\mathrm{d}x = a_0(b-a) + a_1\frac{b^2-a^2}{2} + \dots + a_n\frac{b^{n+1}-a^{n+1}}{n+1}\cdot$$

O altă regulă, evidentă atît din definiția analitică a integralei, cît și din interpretarea ei geometrică, este dată de formula

$$(10)\qquad \int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x.$$

Mai mult, este clar că integrala se anulează dacă $b = a$. Regula de la p. 422

$$(11)\qquad \int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\,\mathrm{d}x,$$

concordă cu ultimele două reguli, *b* deoarece ea corespunde lui (10) pentru $c = a$.

Uneori este convenabil să folosim faptul că valoarea integralei nu depinde de numele particular x ales pentru variabila independentă din $f(x)$; de exemplu,

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b f(u)\,\mathrm{d}u = \int_a^b f(t)\,\mathrm{d}t \text{ etc.}$$

Într-adevăr, o simplă schimbare a numelui coordonatelor sistemului față de care este reprezentat graficul funcției nu modifică aria aflată sub curbă. Aceeași observație se aplică chiar dacă facem anumite modificări în sistemul de coordonate. De exemplu, să deplasăm originea spre dreapta cu o unitate, de la O la O', ca în fig. 265, așa încît x este înlocuită cu o nouă coordo-

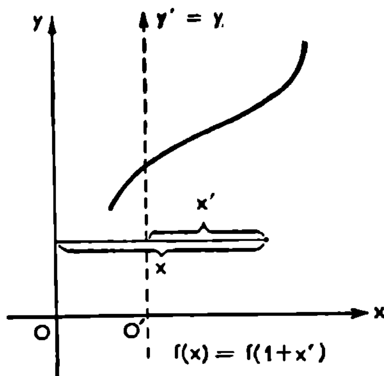


Fig. 265. Deplasarea axei Oy

nată x', legată de prima prin relația $x = 1 + x'$. O curbă de ecuație $y = f(x)$ va avea în noul sistem de coordonate ecuația $y = f(1 + x')$. (De exemplu, $y = 1/x = 1/(1 + x')$.) O arie dată A, aflată sub această curbă, de pildă cea cuprinsă între $x = 1$, și $x = b$, devine în noul sistem de coordonate aria aflată sub arcul cuprins între $x' = 0$ și $x' = b - 1$. Astfel avem

$$\int_1^b f(x)\,dx = \int_0^{b-1} f(1 + x')\,dx'$$

sau, înlocuind pe x' cn u,

$$\int_1^b f(x)\,dx = \int_0^{b-1} f(1 + u)\,du. \tag{12}$$

De exemplu,

$$\int_1^b \frac{1}{x}\,dx = \int_0^{b-1} \frac{1}{1 + u}\,du \tag{12a}$$

şi pentru funcţia $f(x) = x^k$

(12b) $$\int_1^b x^k\, dx = \int_0^{b-1} (1+u)^k\, du.$$

În mod similar,

(12c) $$\int_0^b x^k\, dx = \int_{-1}^{b-1} (1+u)^k du \qquad (k \geqslant 0).$$

Deoarece membrul stîng din (12c) este egal cu $\frac{b^{k+1}}{k+1}$, obţinem

(12d) $$\int_{-1}^{b-1} (1+u)^k\, du = \frac{b^{k+1}}{k+1}.$$

Exerciţii: 1) Calculaţi integrala funcţiei $1 + x + x^2 + \ldots + x^n$ de la 0 la b.

2) Pentru $n > 0$, arătaţi că integrala funcţiei $(1+x)^n$ de la -1 la z este egală cu

$$\frac{(1+z)^{n+1}}{(n+1)}.$$

3) Arătaţi că integrala funcţiei $x^n \sin x$, de la 0 la 1, este mai mică decît $1/(n+1)$. (Indicaţie: Această valoare este integrala funcţiei x^n.)

4) Demonstraţi direct şi cu ajutorul formulei binomului că integrala funcţiei $\frac{(1+x)^n}{n}$, de la -1 la z, este egală cu $\frac{(1+z)^{n+1}}{n(n+1)}$.

În sfîrşit, menţionăm două reguli importante care au forma unor inegalităţi. Aceste reguli permit obţinerea unor evaluări aproximative, dar utile, ale integralelor.

Să presupunem că $b > a$ şi că valorile funcţiei $f(x)$ în acest interval nu depăşesc nicăieri pe acelea ale funcţiei $g(x)$. Atunci avem.

(13) $$\int_a^b f(x) dx \leqslant \int_a^b g(x) dx,$$

după cum rezultă imediat fie din fig. 266, fie din definiţia analitică a integralei. În particular, dacă $g(x) = M$ este o constantă pe care nu o depăşesc valorile funcţiei $f(x)$, avem $\int_a^b g(x)\, dx = \int_a^b M dx = M(b-a)$. Rezultă că

(14) $$\int_a^b f(x)\, dx \leqslant M(b-a).$$

Dacă $f(x)$ nu este negativă, atunci $f(x) = |f(x)|$. Dacă $f(x) < 0$, atunci $|x(f)| > f(x)$. Deci, punînd $g(x) = |f(x)|$ în (13) obținem formula utilă

(15) $$\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b |f(x)|\mathrm{d}x.$$

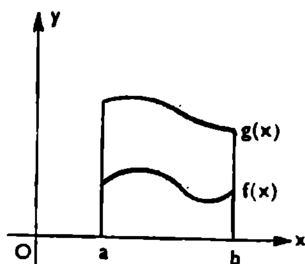


Fig. 266. Compararea integralelor

Deoarece $|-f(x)| = |f(x)|$, avem și

$$-\int_a^b f(x)\mathrm{d}x \leqslant \int_a^b |f(x)|\,\mathrm{d}x,$$

care împreună cu (15) ne dau inegalitatea mai puternică

(16) $$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leqslant \int_a^b |f(x)|\,\mathrm{d}x.$$

§ 2. DERIVATA

1. Derivata ca pantă

În timp ce noțiunea de integrală își are rădăcinile în antichitate, cealaltă noțiune fundamentală a analizei, derivată, a apărut abia în secolul al XVII-lea, datorită lui Fermat și altor matematicieni. Legătura organică dintre aceste două noțiuni, aparent cu totul diferite, descoperită de Newton și Leibniz, a inaugurat o dezvoltare fără precedent a științei matematice.

Fermat era interesat în determinarea maximelor și minimelor unei funcții $y = f(x)$. În graficul unei funcții, un maxim corespunde unui vîrf mai înalt

decît toate celelalte puncte învecinate, în timp ce un minim corespunde unui punct mai coborît decît toate punctele învecinate. În fig. 191, punctul B este un maxim, iar punctul C un minim. Pentru a caracteriza punctele de maxim și minim, este naturală folosirea noțiunii de *tangentă* a unei curbe. Presupunem că graficul nu are colțuri ascuțite sau alte singularități și că, în fiecare punct, el are o direcție determinată, dată de dreapta tangentă în acel punct. În punctele de maxim sau de minim, tangenta la graficul funcției $y = f(x)$ trebuie să fie paralelă cu axa Ox, deoarece, în caz contrar, curba ar urca sau ar coborî în aceste puncte. Această observație sugerează ideea considerării tangentei la curbă în fiecare punct P al graficului funcției $y = f(x)$.

Pentru a caracteriza direcția unei drepte din planul (x, y) se dă de obicei panta ei, care este tangenta trigonometrică a unghiului α, format de sensul pozitiv al axei Ox cu dreapta. Dacă P este un punct al dreptei L, ne deplasăm spre dreapta într-un punct R și apoi în sus sau în jos, pînă în punctul Q de pe dreaptă; atunci panta dreptei $L = \operatorname{tg} \alpha = \frac{RQ}{PR}$. Lungimea PR este luată pozitivă, în timp ce RQ este luată pozitivă sau negativă, după cum dreapta de la R la Q urcă sau coboară, astfel încît panta dă urcarea sau coborîrea pe unitatea de lungime, purtată în lungul orizontalei, atunci cînd ne deplasăm pe dreaptă de la stînga spre dreapta. În fig. 267, panta primei drepte este 2/3, în timp ce panta celei de-a doua este -1.

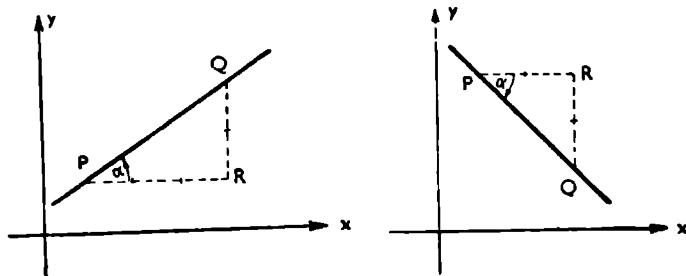


Fig. 267. Pantele dreptelor

Prin panta unei *curbe* într-un punct P înțelegem panta tangentei la curbă în punctul P. Atîta timp cît acceptăm tangenta la o curbă ca noțiune matematică dată intuitiv, rămîne doar problema *găsirii unui procedeu pentru calculul pantei*. Pentru moment vom accepta acest punct de vedere, urmînd să facem în supliment o analiză mai amănunțită a problemelor implicate.

2. Derivata ca limită

Panta unei curbe $y = f(x)$ în punctul $P(x, y)$ nu poate fi calculată referindu-ne doar la punctul P al curbei. În schimb, trebuie să recurgem la un proces de trecere la limită, foarte asemănător cu acela care a intervenit în calculul ariei aflate sub o curbă. Acest proces de trecere la limită constituie baza calculului diferențial. Să considerăm pe curbă un alt punct P_1, vecin cu P, de coordonate x_1, y_1. Să notăm cu t_1 dreapta care unește punctul P cu P_1; ea este o secantă a curbei, care aproximează tangenta în P, atunci cînd P_1 este apropiat de P. Să notăm unghiul format de sensul pozitiv al axei Ox cu t_1 prin α_1. Dacă facem ca x_1 să se apropie de x, punctul P_1 se va deplasa de-a lungul curbei spre P, iar secanta t_1 va avea ca poziție limită *tangenta* t la curbă în punctul P. Dacă α este unghiul format de sensul pozitiv al axei Ox cu t, atunci cînd $x_1 \to x$*,

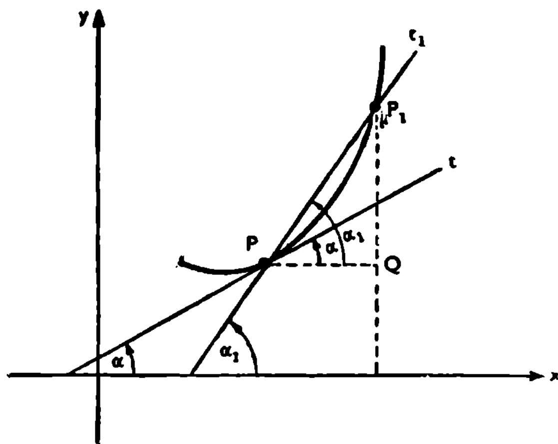


Fig. 268. Derivata ca limită

$$y_1 \to y,\ P_1 \to P,\ t_1 \to t \text{ și } \alpha_1 \to \alpha.$$

Tangenta este limita secantei, iar panta tangentei este limita pantei secantei.

Deși nu avem nici o expresie explicită pentru panta tangentei t, panta secantei t_1 este dată de formula

$$\text{panta lui } t_1 = \frac{y_1 - y}{x_1 - x} = \frac{f(x_1) - f(x)}{x_1 - x},$$

* Notația folosită aici este puțin diferită de aceea din cap. VI, deoarece acolo avem $x \to x_1$, ultima valoare fiind cea fixată. Nu poate să apară nici o confuzie din această permutare a simbolurilor.

sau, dacă notăm din nou cu Δ operația de formare a unei diferențe,

$$\text{panta lui } t_1 = \frac{\Delta y}{\Delta x} = \frac{\Delta f(x)}{\Delta x}.$$

Panta secantei t_1 este un „cît al creșterilor” — creșterea Δy a valorilor funcției, împărțită la creșterea Δx a valorilor variabilei independente. Mai mult,

$$\text{panta lui } t = \text{limita pantei lui } t_1 = \lim \frac{f(x_1) - f(x)}{x_1 - x} = \lim \frac{\Delta y}{\Delta x}, \text{ unde}$$

limitele sînt calculate pentru $x_1 \to x$, adică atunci cînd $\Delta x = x_1 - x \to 0$. *Panta tangentei* t *la curbă este limita cîtului creșterilor* $\Delta y/\Delta x$, *atunci cînd* $\Delta x = x_1 - x$ *tinde spre zero.*

Funcția inițială $f(x)$ dădea înălțimea curbei $y = f(x)$, corespunzătoare valorii x. Acum putem considera panta curbei pentru un punct P variabil, de coordonate x și $y[= f(x)]$ ca o nouă funcție de x, pe care o notăm cu $f'(x)$ și o denumim *derivata* funcției $f(x)$. Procesul de trecere la limită prin care ea se obține se numește *derivarea* funcției $f(x)$. Acest proces este o operație care atașează unei funcții date $f(x)$ o altă funcție $f'(x)$, potrivit unei anumite reguli, tot așa cum funcția $f(x)$ este definită printr-o regulă care asociază oricărei valori a variabilei x valoarea $f(x)$:

$$f(x) = \text{înălțimea curbei } y = f(x) \text{ în punctul } x,$$
$$f'(x) = \text{panta curbei } y = f(x) \text{ în punctul } x.$$

Cuvîntul „derivare”[1] provine din faptul că $f'(x)$ este limita diferenței $f(x_1) - f(x)$ împărțită la diferența $x_1 - x$:

$$(1) \qquad f'(x) = \lim \frac{f(x_1) - f(x)}{x_1 - x,} \text{ cînd } x_1 \to x.$$

O altă notație, folosită uneori, este:

$$f'(x) = Df(x),$$

litera „D” fiind o prescurtare pentru „derivata lui”; o altă notație pentru derivata lui $y = f(x)$ este aceea a lui Leibniz:

$$\frac{dy}{dx} \text{ sau } \frac{df(x)}{dx},$$

pe care o vom discuta în § 4 și care indică natura derivatei ca limită a cîtului creșterilor $\Delta y/\Delta x$ sau $\Delta f(x)/\Delta x$.

[1] În original este folosit cuvîntul *differentiation*, care se potrivește mai bine cu explicația dată. — *N.T.*

Dacă descriem curba $y = f(x)$ în sensul valorilor crescătoare ale lui x, atunci o *derivată pozitivă* $f'(x) > 0$ într-un punct înseamnă o *curbă crescătoare* (valori crescătoare ale lui y), iar o *derivată negativă*, $f'(x) < 0$, înseamnă o *curbă descrescătoare*, în timp ce $f'(x) = 0$ înseamnă o direcţie orizontală a curbei pentru

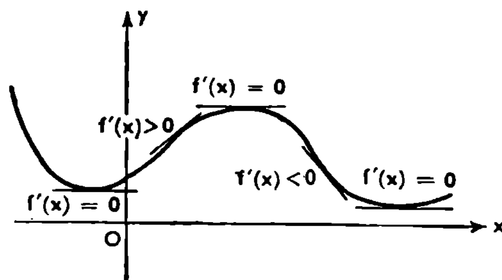


Fig. 269. Semnul derivatei

valoarea corespunzătoare a lui x. Într-un maxim sau minim, panta trebuie să fie nulă (fig. 269). Deci rezolvînd ecuaţia $f'(x) = 0$ în raport cu x, putem găsi punctele de maxim şi de minim, aşa cum a făcut pentru prima oară Fermat.

3. Exemple

Consideraţiile care duc la definiţia (1) ar putea să pară lipsite de valoare practică. O problemă a fost înlocuită cu alta: în loc de a cere panta tangentei la o curbă $y = f(x)$ într-un punct, se cere evaluarea unei limite, (1), care la prima vedere pare o problemă tot atît de dificilă. Însă îndată ce părăsim domeniul generalităţilor şi considerăm funcţii particulare $f(x)$, vom obţine rezultate palpabile.

Cea mai simplă funcţie de acest fel este $f(x) = c$, unde c este o constantă. Graficul funcţiei $y = f(x) = c$ este o dreaptă orizontală, care coincide cu toate tangentele ei şi este evident că

$$f'(x) = 0$$

pentru toate valorile lui x. Aceasta rezultă din definiţia (1), pentru că

$$\frac{\Delta y}{\Delta x} = \frac{f(x_1) - f(x)}{x_1 - x} = \frac{c - c}{x_1 - x} = \frac{0}{x_1 - x} = 0,$$

astfel încît

$$\lim \frac{f(x_1) - f(x)}{x_1 - x} = 0, \text{ cînd } x_1 \to x.$$

Apoi să considerăm funcția simplă $y = f(x) = x$, al cărei grafic este prima bisectoare. Din punct de vedere geometric este clar că

$$f'(x) = 1$$

pentru toate valorile lui x, iar definiția analitică (1) ne dă din nou

$$\frac{f(x_1) - f(x)}{x_1 - x} = \frac{x_1 - x}{x_1 - x} = 1,$$

astfel încît

$$\lim \frac{f(x_1) - f(x)}{x_1 - x} = 1, \text{ cînd } x_1 \to x.$$

Cel mai simplu exemplu nebanal este derivarea funcției

$$y = f(x) = x^2,$$

ceea ce înseamnă găsirea pantei unei parabole. Aceasta este cel mai simplu caz care ne învață cum trebuie efectuată trecerea la limită, atunci cînd rezultatul nu este evident de la început. Avem:

$$\frac{\Delta y}{\Delta x} = \frac{f(x_1) - f(x)}{x_1 - x} = \frac{x_1^2 - x^2}{x_1 - x}.$$

Dacă am încerca să trecem la limită direct în numărător și numitor, am obține expresia lipsită de sens 0/0. Putem evita însă acest impas, scriind din nou cîtul creșterilor și simplificînd factorul $x_1 - x$ care ne incomodează *înainte de a trece la limită*. (La evaluarea limitei cîtului creșterilor, considerăm doar valorile $x_1 \neq x$, astfel încît acest lucru este permis; conform p. 324.) Obținem astfel expresia:

$$\frac{x_1^2 - x^2}{x_1 - x} = \frac{(x_1 - x)(x_1 + x)}{x_1 - x} = x_1 + x.$$

După simplificare, aflarea limitei cînd $x_1 \to x$ nu mai prezintă nici o dificultate. Limita se obține „prin substituție“, pentru că noua formă $x_1 + x$ a cîtului creșterilor este continuă și limita unei funcții continue, cînd $x_1 \to x$ este pur și simplu valoarea funcției pentru $x_1 = x$; în cazul nostru, $x + x = 2x$, astfel încît

$$f'(x) = 2x \quad \text{pentru } f(x) = x^2.$$

În mod analog putem demonstra că pentru $f(x) = x^3$ avem $f'(x) = 3x^2$. Într-adevăr, cîtul creşterilor

$$\frac{\Delta y}{\Delta x} = \frac{f(x_1) - f(x)}{x_1 - x} = \frac{x_1^3 - x^3}{x_1 - x}$$

poate fi simplificat cu ajutorul formulei $x_1^3 - x^3 = (x_1 - x)(x_1^2 + x_1 x + x^2)$; numitorul $\Delta x = x_1 - x$ se simplifică şi obţinem expresia continuă

$$\frac{\Delta y}{\Delta x} = x_1^2 + x_1 x + x^2.$$

Dacă facem acum pe x_1 să tindă către x, această expresie tinde spre $x^2 + x^2 + x^2$ şi obţinem la limită $f'(x) = 3x^2$. În general, pentru

$$f(x) = x^n,$$

unde n este un întreg pozitiv oarecare, obţinem derivata

$$f'(x) = nx^{n-1}.$$

Exerciţiu: Demonstraţi acest rezultat. (Folosiţi formula algebrică

$$x_1^n - x^n = (x_1 - x)(x_1^{n-1} + x_1^{n-2}x + x_1^{n-3}x^2 + \ldots + x_1 x^{n-2} + x^{n-1}).)$$

Un alt exemplu de artificii simple, care permit determinarea explicită a derivatei, este dat de funcţia

$$y = f(x) = \frac{1}{x}.$$

Avem

$$\frac{\Delta y}{\Delta x} = \frac{y_1 - y}{x_1 - x} = \left(\frac{1}{x_1} - \frac{1}{x}\right) \cdot \frac{1}{x_1 - x} = \frac{x - x_1}{x_1 x} \cdot \frac{1}{x_1 - x}.$$

Putem simplifica din nou şi găsim $\frac{\Delta y}{\Delta x} = -\frac{1}{x_1 x}$, care este continuă în $x_1 = x$; deci la limită avem

$$f'(x) = -\frac{1}{x^2}.$$

Desigur, nici derivata şi nici funcţia nu sînt definite pentru $x = 0$.

Exerciţii: Demonstraţi în mod asemănător că pentru $f(x) = 1/x^2$, $f'(x) = -2/x^3$; pentru $f(x) = 1/x^n$, $f'(x) = -n/x^{n+1}$ pentru $f(x) = (1 + x)^n$ avem $f'(x) = n(1 + x)^{n-1}$.

Vom calcula acum derivata funcţiei

$$y = f(x) = \sqrt{x}.$$

Pentru cîtul creşterilor obţinem expresia:

$$\frac{y_1 - y}{x_1 - x} = \frac{\sqrt{x_1} - \sqrt{x}}{x_1 - x}.$$

Ţinînd seama de formula $x_1 - x = (\sqrt{x_1} - \sqrt{x})(\sqrt{x_1} + \sqrt{x})$, putem simplifica şi obţinem expresia continuă:

$$\frac{y_1 - y}{x_1 - x} = \frac{1}{\sqrt{x_1} + \sqrt{x}}.$$

Prin trecere la limită obţinem

$$f'(x) = \frac{1}{2\sqrt{x}}.$$

Exerciţii: Demonstraţi că pentru $f(x) = \frac{1}{\sqrt{x}}$ avem $f'(x) = -\frac{1}{2(\sqrt{x})^3}$; pentru $(x) = \sqrt[3]{x}$ avem $f'(x) = \frac{1}{3\sqrt[3]{x^2}}$; pentru $f(x) = \sqrt{1 - x^3}$ avem $f'(x) = \frac{-x}{\sqrt{1 - x^2}}$; pentru $f(x) = \sqrt[n]{x}$ avem $f'(x) = \frac{1}{n\sqrt[n]{x^{n-1}}}$.

4. Derivatele funcţiilor trigonometrice

Vom trata acum problema importantă a *derivării funcţiilor trigonometrice*. În cele ce urmează vom folosi ca unitate de măsură a unghiurilor radianul.

Pentru a deriva funcţia $y = f(x) = \sin x$ punem $x_1 - x = h$, astfel încît $x_1 = x + h$ şi $f(x_1) = \sin x_1 = \sin (x + h)$. Din formula trigonometrică cunoscută pentru $\sin (A + B)$, deducem

$$f(x_1) = \sin (x + h) = \sin x \cos h + \cos x \sin h.$$

Deci

$$(2) \quad \frac{f(x_1) - f(x)}{x_1 - x} = \frac{\sin (x + h) - \sin x}{h} = \cos x\left(\frac{\sin h}{h}\right) + \sin x\left(\frac{\cos h - 1}{h}\right).$$

Dacă facem acum ca x_1 să tindă către x, atunci h tinde spre 0, sin h spre 0, iar cos h spre 1. Mai mult, din rezultatele de la p. 324 avem

$$\lim \frac{\sin h}{h} = 1$$

şi

$$\lim \frac{\cos h - 1}{h} = 0.$$

Prin urmare, membrul drept din (2) tinde spre cos x, deci am obţinut rezultatul:

Funcţia **f(x) = sin x *are derivata* f′(x) = cos x sau, pe scurt,**

$$D \sin x = \cos x.$$

Exerciţiu: Demonstraţi că $D \cos x = -\sin x$.

Pentru a deriva funcţia $f(x) = \text{tg } x$, scriem $\text{tg } x = \sin x/\cos x$ şi obţinem:

$$\frac{f(x+h) - f(x)}{h} = \left(\frac{\sin(x+h)}{\cos(x+h)} - \frac{\sin x}{\cos x}\right) \cdot \frac{1}{h} =$$

$$= \frac{\sin(x+h)\cos x - \cos(x+h)\sin x}{h} \cdot \frac{1}{\cos(x+h)\cos x} =$$

$$= \frac{\sin h}{h} \cdot \frac{1}{\cos(x+h)\cos x}.$$

(Ultima egalitate rezultă din formula $\sin(A - B) = \sin A \cos B - \cos A \sin B$, cu $A = x + h$ şi $B = h$.) Dacă facem acum pe h să tindă spre zero, atunci $(\sin h)/h$ tinde spre 1, $\cos(x + h)$ tinde spre $\cos x$ şi deducem:

Derivata funcţiei **f(x) = tg x *este* f′(x) = 1/cos²x, sau** $D \text{ tg } x = \dfrac{1}{\cos^2 x}$.

Exerciţiu: Arătaţi că $D \text{ ctg } x = -\dfrac{1}{\sin^2 x}$.

*5 Derivabilitatea şi continuitatea

Derivabilitatea unei funcţii implică continuitatea ei. Într-adevăr, dacă limita cîtului $\Delta y/\Delta x$ există cînd Δx tinde către zero, se observă uşor că creşterea Δy a funcţiei $f(x)$ trebuie să devină arbitrar de mică, atunci cînd diferenţa Δx tinde spre zero. Deci, ori de cîte ori putem deriva o funcţie, con-

tinuitatea ei este automat asigurată; de aceea, ne vom lipsi de menţionarea sau demonstrarea explicită a continuităţii funcţiilor derivabile care intervin în acest capitol, afară de cazul în care avem motive speciale să o facem.

6. Derivata şi viteza. Derivata de ordinul doi şi acceleraţia

Discuţia precedentă asupra derivatei a fost efectuată în legătură cu noţiunea geometrică de grafic al unei funcţii. Dar însemnătatea noţiunii de derivată nu este cîtuşi de puţin limitată la problema găsirii pantei tangentei la o curbă. Şi mai importantă pentru ştiinţele naturii este problema calculării *vitezei de variaţie* a unei cantităţi $f(t)$, care depinde de timpul t. Acesta este punctul de vedere din care Newton a abordat calculul diferenţial. Newton dorea în particular să analizeze fenomenul vitezei, în care timpul şi poziţia unei particule mobile sînt considerate elemente variabile sau, după expresia lui Newton, „cantităţi fluente".

Dacă o particulă se deplasează de-a lungul unei drepte, axa Ox, mişcarea ei este complet determinată prin cunoaşterea poziţiei x în orice moment t, ca o funcţie $x = f(t)$. O „mişcare uniformă", cu viteză constantă b, de-a lungul axei Ox, este definită printr-o funcţie liniară $x = a + bt$, unde a este coordonata particulei la momentul $t = 0$.

În plan, mişcarea unei particule este descrisă de două funcţii

$$x = f(t), \quad y = g(t),$$

care caracterizează cele două coordonate ca funcţii de timp. În particular, o mişcare uniformă corespunde unei perechi de funcţii liniare,

$$x = a + bt, \quad y = c + dt,$$

unde b şi d sînt cele două „componente" ale vitezei constante iar a şi c sînt coordonatele particulei în momentul $t = 0$; traiectoria particulei este o dreaptă de ecuaţie $(x - a)d - (y - c)b = 0$, care se obţine prin eliminarea timpului t între cele două ecuaţii de mai sus.

Dacă o particulă se deplasează în planul vertical (x, y) sub influenţa gravitaţiei, atunci, aşa cum se arată în fizica elementară, mişcarea ei este descrisă de două ecuaţii:

$$x = a + bt, \quad y = c + dt - \frac{1}{2}gt^2,$$

unde a, b, c, d sînt constante care depind de starea iniţială a particulei, iar g este acceleraţia gravitaţiei, egală aproximativ cu 981, dacă timpul se

măsoară în secunde şi distanţa în centimetri. Traiectoria particulei, obţinută prin eliminarea lui t între cele două ecuaţii, este parabola

$$y = c + \frac{d}{b}(x - a) - \frac{1}{2} g \frac{(x-a)^2}{b^2},$$

dacă $b \neq 0$; în caz contrar, ea este o parte a axei verticale.

Dacă o particulă este constrînsă la o deplasare de-a lungul unei curbe din plan (ca de pildă, un tren de-a lungul şinelor), atunci mişcarea lui poate fi descrisă dînd lungimea arcului s, măsurat de-a lungul curbei dintr-un punct iniţial fixat P_0 pînă în poziţia P a particulei la momentul t, ca funcţie de t; $s = f(t)$. De exemplu, pe cercul unitate, $x^2 + y^2 = 1$, funcţia $s = ct$ descrie o rotaţie uniformă cu viteza c de-a lungul cercului.

Exerciţii: *Trasaţi traiectoriile mişcării plane descrise de

1) $x = \sin t$, $y = \cos t$. 2) $x = \sin 2t$, $y = \sin 3t$. 3) $x = \sin 2t$, $y = 2 \sin 3t$, 4) În mişcarea parabolică descrisă mai sus să presupunem că particula se află în origine la momentul $t = 0$ şi că $b > 0$. $d > 0$. Găsiţi coordonatele celui mai înalt punct al traiectoriei. Găsiţi momentul t şi valoarea lui x pentru a doua intersecţie a traiectoriei cu axa Ox.

Primul scop al lui Newton a fost determinarea vitezei unei mişcări neuniforme. Pentru simplificare, să considerăm mişcarea unei particule de-a lungul unei drepte, dată de o funcţie $x = f(t)$. Dacă mişcarea ar fi uniformă, adică viteza este constantă, atunci această viteză ar putea fi găsită luînd două valori t şi t_1, ale timpului cu valorile corespunzătoare $x = f(t)$ şi $x_1 = f(t_1)$ ale poziţiei, şi formînd cîtul

$$(2) \qquad v = \text{viteza} = \frac{\text{distanţa}}{\text{timp}} = \frac{x_1 - x}{t_1 - t} = \frac{f(t_1) - f(t)}{t_1 - t}.$$

De exemplu, dacă t se măsoară în ore şi x în mile, atunci pentru $t_1 - t = 1$, $x_1 - x$ va fi numărul de mile parcurs într-o oră iar v va fi viteza în mile pe oră. Afirmaţia că viteza mişcării este constantă înseamnă, pur şi simplu, că cîtul creşterilor

$$(3) \qquad \frac{f(t_1) - f(t)}{t_1 - t}$$

este acelaşi pentru toate valorile lui t şi t_1. Dar atunci cînd mişcarea nu este uniformă, ca în cazul unui corp aflat în cădere liberă, a cărui viteză creşte pe măsură ce el cade, cîtul (3) nu mai dă viteza *la momentul t*, ci doar *viteza medie* în intervalul de timp cuprins între t şi t_1. Pentru a obţine viteza la

momentul t, trebuie să luăm limita vitezei medii, cînd t_1 tinde spre t. Astfel definim, aşa cum a făcut Newton,

$$\text{(4)} \qquad \text{viteza la momentul } t = \lim \frac{f(t_1) - f(t)}{t_1 - t} = f'(t).$$

Cu alte cuvinte, viteza este derivata coordonatei distanţei în raport cu timpul sau ,,viteza instantanee de variaţie" a distanţei în raport cu timpul (spre deosebire de viteza *medie* de variaţie, dată de (3)).

Viteza de variaţie a vitezei se numeşte *acceleraţie.* Ea este pur şi simplu derivata derivatei, care se notează de obicei cu $f''(t)$ şi se numeşte derivata de ordinul doi a lui $f(t)$.

Galilei a observat că pentru un corp în cădere liberă, distanţa verticală x, cu care el cade în intervalul de timp t, este dată de formula

$$\text{(5)} \qquad x = f(t) = \frac{1}{2} g t^2,$$

unde g este constanta gravitaţională. Derivînd funcţia din (5), rezultă că viteza v a corpului la momentul t este dată de

$$\text{(6)} \qquad v = f'(t) = gt,$$

iar acceleraţia α este dată de

$$\alpha = f''(t) = g,$$

care este constantă.

Să presupunem că trebuie aflată viteza corpului după 2 secunde de cădere. Viteza *medie* în intervalul de timp de la $t = 2$ la $t = 2{,}1$ este

$$\frac{\frac{1}{2} g \cdot (2{,}1)^2 - \frac{1}{2} g \cdot (2)^2}{2{,}1 - 2} = 2011{,}05 \text{ cm/s}.$$

Substituind $t = 2$ în (6), obţinem viteza *instantanee* după două secunde, egală cu

$$v = 1962 \text{ cm/s}.$$

Exerciţiu: Care este viteza medie a corpului în intervalul de timp de la $t = 2$ la $t = 2{,}01$? Dar de la $t = 2$ la $t = 2{,}001$?

Pentru mişcarea plană, cele două derivate $f'(t)$ şi $g'(t)$ ale funcţiilor $x = f(t)$ şi $y = g(t)$ definesc componentele vitezei. Pentru mişcarea efectuată de-a lungul unei curbe fixate, viteza va fi definită de derivata funcţiei $s = f(t)$, unde s este lungimea arcului.

7. Interpretarea geometrică a derivatei de ordinul doi

Derivata de ordinul doi este importantă și în analiză și geometrie pentru că $f''(x)$, exprimînd viteza de variație a pantei $f'(x)$ a curbei $y = f(x)$, dă o indicație asupra modului în care este încovoiată curba. Dacă $f''(x)$ este pozitivă într-un interval, atunci viteza de variație a lui $f'(x)$ este pozitivă. O viteză

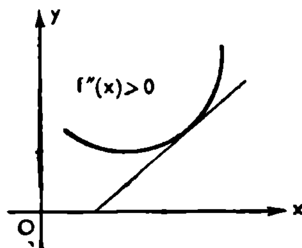


Fig. 270.

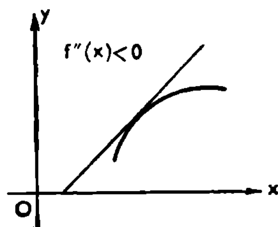


Fig. 271.

pozitivă de variație a unei funcții indică faptul că valorile funcției cresc o dată cu x. De aceea $f''(x) > 0$ înseamnă că panta $f'(x)$ crește o dată cu x, astfel încît curba devine mai abruptă, acolo unde ea are o pantă pozitivă și mai puțin abruptă, acolo unde are o pantă negativă. Spunem că curba este *convexă* sau cu *concavitatea în sus* (fig. 270).

În mod similar, dacă $f''(x) < 0$, curba $y = f(x)$ este *concavă* sau cu *concavitatea în jos* (fig. 271).

Parabola $y = f(x) = x^2$ are concavitatea în sus peste tot, pentru că $f''(x) = 2$ este pozitivă întotdeauna. Curba $y = f(x) = x^3$ are concavitatea în sus pentru $x > 0$ și concavitatea în jos pentru $x < 0$ (fig. 153), pentru că $f''(x) = 6x$, după cum cititorul poate arăta cu ușurință. Observăm totodată că pentru $x = 0$ avem $f'(x) = 3x^2 = 0$ (însă nu avem maxim sau minim); de asemenea, $f''(x) = 0$, pentru $x = 0$. Acest punct se numește *punct de inflexiune*. Într-un astfel de punct, tangenta, în cazul nostru axa Ox, traversează curba.

Dacă s este lungimea arcului de curbă, iar α unghiul făcut de tangentă cu axa Ox, atunci $\alpha = h(s)$ va fi o funcție de s. Cînd parcurgem curba, $\alpha = h(s)$ va varia. Viteza de variație $h'(s)$ se numește *curbura* curbei în punctul în care lungimea arcului este egală cu s. Menționăm, fără demonstrație, faptul că curbura k poate fi exprimată în funcție de derivatele de ordinul întîi și doi ale funcției $y = f(x)$ care definește curba, prin formula

$$k = f''(x)/(1 + f'(x)^2)^{3/2}.$$

8. Maxime și minime

Putem găsi maximele și minimele unei funcții date $f(x)$ formînd mai întîi derivata $f'(x)$, apoi găsind valorile în care se anulează această derivată și cercetînd care dintre aceste valori dau maxime și care minime. Ultima problemă

poate fi decisă cu ajutorul derivatei de ordinul doi $f''(x)$, al cărei semn indică forma convexă sau concavă a graficului şi a cărei anulare indică, de obicei, un punct de inflexiune, în care nu are loc nici un extrem. Observînd semnele lui $f'(x)$, şi $f''(x)$, putem nu numai să determinăm extremele, dar şi forma graficului funcţiei. Această metodă ne dă valorile lui x, în care au loc extremele ; pentru a găsi valorile corespunzătoare ale lui $y = f(x)$, trebuie să substituim aceste valori ale lui x în $f(x)$.

De exemplu, să considerăm polinomul

$$f(x) = 2x^3 - 9x^2 + 12x + 1$$

şi obţinem

$$f'(x) = 6x^2 - 18x + 12, \quad f''(x) = 12x - 18.$$

Rădăcinile ecuaţiei pătratice $f'(x) = 0$ sînt $x_1 = 1$, $x_2 = 2$ şi avem $f''(x_1) = -6 < 0$, $f''(x_2) = 6 > 0$. Deci $f(x)$ are un maxim, $f(x_1) = 6$ şi un minim, $f(x_2) = 5$.

Exerciţii : 1) Schiţaţi graficul funcţiei considerate mai sus.

2) Discutaţi şi schiţaţi graficul funcţiei $f(x) = (x^2 - 1)(x^2 - 4)$.

3) Găsiţi minimul funcţiilor $x + 1/x$, $x + a^2/x$, $px + q/x$, unde p şi q sînt numere pozitive. Aceste funcţii au maxim?

4) Găsiţi maximele şi minimele funcţiilor $\sin x$ şi $\sin (x^2)$.

§ 3. TEHNICA DERIVĂRII

Pînă acum eforturile noastre s-au îndreptat spre derivarea cîtorva funcţii particulare, prin transformarea cîturilor creşterilor, înainte de a face trecerea la limită. Un pas decisiv s-a făcut atunci cînd, datorită lucrărilor lui Leibniz, Newton şi ale continuatorilor lor, artificiile particulare au fost înlocuite prin metode generale puternice. Cu ajutorul acestor metode putem deriva aproape automat orice funcţie care se întîlneşte în mod normal în matematică, cu condiţia să stăpînim cîteva reguli simple şi să recunoaştem momentul aplicării lor. Astfel derivarea a căpătat trăsătura unui „algoritm" de calcul şi tocmai acesta este aspectul teoriei exprimat prin termenul „calcul".

Nu putem insista mai mult asupra amănuntelor acestei tehnici. Vor fi menţionate doar cîteva reguli simple.

a) *Derivarea unei sume.* Dacă a şi b sînt constante, iar funcţia $k(x)$ este dată de

$$k(x) = af(x) + bg(x),$$

atunci, după cum cititorul poate verifica cu uşurinţă,

$$k'(x) = af'(x) + bg'(x).$$

O regulă asemănătoare este valabilă pentru orice număr de termeni.

b) *Derivarea unui produs.* Pentru un produs,

$$p(x) = f(x)g(x),$$

derivata este

$$p'(x) = f(x)g'(x) + g(x)f'(x).$$

Acest lucru se demonstrează cu uşurinţă, folosind următorul artificiu : adunînd şi scăzînd acelaşi termen, scriem

$$p(x + h) - p(x) = f(x + h)g(x + h) - f(x)g(x) =$$

$$= f(x + h)g(x + h) - f(x + h)\,g(x) + f(x + h)g(x) - f(x)g(x)$$

şi obţinem, combinînd primii doi şi ultimii doi termeni,

$$\frac{p(x + h) - p(x)}{h} = f(x + h)\frac{g(x + h) - g(x)}{h} + g(x)\frac{f(x + h) - f(x)}{h}.$$

Acum să facem pe h să tindă spre 0 ; deoarece $f(x + h)$ tinde spre $f(x)$, afirmaţia rezultă imediat.

Exerciţiu : Demonstraţi că funcţia $p(x) = x^n$ are derivata $p'(x) = nx^{n-1}$. (Indicaţie : scrieţi $x^n = x \cdot x^{n-1}$ şi folosiţi inducţia matematică.

Folosind regulile (a) şi (b), putem deriva orice polinom :

$$f(x) = a_0 + a_1x + \ldots + a_nx^n,$$

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + \ldots + na_nx^{n-1}.$$

Ca aplicaţie, putem demonstra *teorema binomului* (cf. p. 32). Această teoremă se referă la dezvoltarea lui $(1 + x)^n$, ca polinom :

(1) $$f(x) = (1 + x)^n = 1 + a_1x + a_2x^2 + a_3x^3 + \ldots + a_nx^n,$$

şi afirmă că coeficientul a_k este dat de formula

(2) $$a_k = \frac{n(n - 1) \ldots (n - k + 1)}{k\,!}.$$

Desigur, $a_n = 1$.

Am văzut (exerciţiul de la p. 438) că membrul stîng al egalităţii (1) dă prin derivare $n(1 + x)^{n-1}$. Prin urmare, în baza rezultatului precedent, obţinem

(3) $$n(1 + x)^{n-1} = a_1 + 2a_2x + 3a_3x^2 + \ldots + na_nx^{n-1}.$$

În această formulă să punem acum $x = 0$ şi găsim că $n = a_1$, care este tocmai formula (2) pentru $k = 1$. Apoi derivînd din nou egalitatea (3), obţinem

$$n(n-1)(1+x)^{n-2} = 2a_2 + 3 \cdot 2a_3x + \ldots + n(n-1)a_nx^{n-2}.$$

Punînd $x = 0$, găsim că $n(n-1) = 2a_2$, conform formulei (2) pentru $k = 2$.

Exerciţiu: Demonstraţi formula (2) pentru $k = 3$, 4 şi pentru un k oarecare, prin inducţie matematică.

c) *Derivarea unui cît.* Dacă

$$q(x) = \frac{f(x)}{g(x)},$$

atunci

$$g'(x) = \frac{g(x)f'(x) - f(x)g'(x)}{(g(x))^2}.$$

Demonstraţia este lăsată ca exerciţiu. (Desigur, trebuie să presupunem că $g(x) \neq 0$.)

Exerciţiu: Deduceţi din această regulă formulele de la p. 440, pentru derivarea funcţiilor tg x şi ctg x, folosind pe acelea pentru derivarea funcţiilor sin x şi cos x. Demonstraţi că derivatele funcţiilor sec $x = 1/\cos x$ şi cosec $x = 1/\sin x$ sînt $\sin x/\cos^3 x$ şi respectiv $-\cos x/\sin^2 x$.

Aşadar, acum sîntem în stare să derivăm orice funcţie care poate fi scrisă sub forma unui cît de două polinoame. De exemplu,

$$f(x) = \frac{1-x}{1+x}$$

are derivata

$$f'(x) = \frac{-(1+x)-(1-x)}{(1+x)^2} = -\frac{2}{(1+x)^2}.$$

Exerciţiu: Derivaţi funcţia

$$f(x) = \frac{1}{x^m} = x^{-m},$$

unde m este un întreg pozitiv. Rezultatul este

$$f'(x) = -mx^{-m-1}.$$

d) *Derivarea funcțiilor inverse.* Dacă

$$y = f(x) \quad \text{şi} \quad x = g(y)$$

sînt funcții inverse (de exemplu, $y = x^2$ şi $x = \sqrt{y}$), atunci derivatele lor sînt inverse una alteia:

$$g'(y) = \frac{1}{f'(x)} \text{ sau } Dg(y) \cdot Df(x) = 1.$$

Acest lucru se demonstrează cu uşurință întorcîndu-ne respectiv la cîturile creşterilor $\frac{\Delta y}{\Delta x}$ şi $\frac{\Delta x}{\Delta y}$, inverse unul altuia; el poate fi observat şi din interpretarea geometrică a funcției inverse, dată la p. 298, dacă considerăm panta tangentei față de axa Oy, în locul pantei față de axa Ox.

Ca exemplu, să derivăm funcția

$$y = f(x) = \sqrt[m]{x} = x^{\frac{1}{m}}$$

inversă funcției $x = y^m$. (A se vedea, de asemenea, metoda mai directă de la p. 438, pentru $m = 1/2$.) Deoarece ultima funcție are ca derivată expresia my^{m-1}, avem

$$f'(x) = \frac{1}{my^{m-1}} = \frac{1}{m} \cdot \frac{y}{y^m} = \frac{1}{m} y\, y^{-m},$$

de unde, făcînd substituțiile $y = x^{1/m}$ şi $y^{-m} = x^{-1}$, găsim $f'(x) = \frac{1}{m} x^{\frac{1}{m}-1}$

sau

$$D(x^{1/m}) = \frac{1}{m} x^{\frac{1}{m}-1}.$$

Un alt exemplu: să derivăm *funcția trigonometrică inversă* (cf. p. 298):

$$y = \text{arc tg } x,$$

ceea ce înseamnă acelaşi lucru cu

$$x = \text{tg } y.$$

Aici variabila y, care este măsurată în radiani, este restrînsă la intervalul $-\frac{\pi}{2} < y < \frac{\pi}{2}$, pentru a asigura univocitatea funcției inverse.

Deoarece avem (cf. p. 440) $D\,\text{tg}\,y = 1/\cos^2 y$ şi deoarece $1/\cos^2 y = (\sin^2 y + \cos^2 y)/\cos^2 y = 1 + \text{tg}^2 y = 1 + x^2$, găsim

$$D\,\text{arc tg}\ x = \frac{1}{1 + x^2}\cdot$$

În acelaşi mod, cititorul poate deduce formulele

$$D\,\text{arc ctg}\ x = -\frac{1}{1 + x^2}$$

$$D\,\text{arc sin}\ x = \frac{1}{\sqrt{1 - x^2}}$$

$$D\,\text{arc cos}\ x = -\frac{1}{\sqrt{1 - x^2}}\cdot$$

În sfîrşit, ajungem la regula importantă pentru

e) *Derivarea funcţiilor compuse.* Astfel de funcţii sînt compuse din două (sau mai multe) funcţii mai simple (cf. p. 299). De exemplu, $z = \sin(\sqrt{x})$ este compusă din $z = \sin y$ şi din $y = \sqrt{x}$; funcţia $z = \sqrt{x} + \sqrt{x^5}$ este compusă din $z = y + y^5$ şi din $y = \sqrt{x}$; $z = \sin(x^2)$ este compusă din $z = \sin y$ şi din $y = x^2$; $z = \sin(1/x)$ este compusă din $z = \sin y$ şi din $y = 1/x$.

Dacă sînt date două funcţii,

$$z = g(y) \quad \text{şi} \quad y = f(x),$$

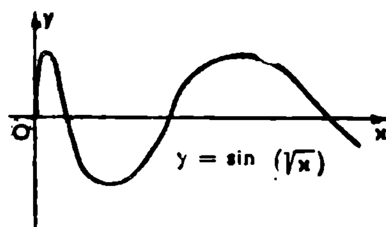


Fig. 272.

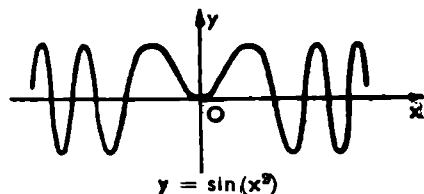


Fig. 273.

şi dacă a doua este substituită în prima, obţinem funcţia compusă

$$z = k(x) = g[f(x)].$$

Afirmăm că

(4) $$k'(x) = g'(y)f'(x).$$

Într-adevăr, dacă scriem

$$\frac{k(x_1) - k(x)}{x_1 - x} = \frac{z_1 - z}{y_1 - y} \cdot \frac{y_1 - y}{x_1 - x},$$

unde $y_1 = f(x_1)$ şi $z_1 = g(y_1) = k(x_1)$ şi dacă facem apoi pe x_1 să tindă spre x, membrul stîng tinde spre $k'(x)$ şi cei doi factori din membrul drept tind respectiv spre $g'(y)$ şi $f'(x)$, ceea ce demonstrează formula (4).

În această demonstraţie am presupus că $y_1 - y \neq 0$, pentru că am împărţit cu $\Delta y = y_1 - y$ şi nu putem folosi valorile x_1, pentru care $y_1 - y = 0$. Formula (4) rămîne însă valabilă şi dacă $\Delta y = 0$ într-un interval din jurul lui x; y este atunci constantă, $f'(x) = 0$, $k(x) = g(y)$ este constantă în raport cu x (deoarece y nu variază o dată cu x) şi deci $k'(x) = 0$, aşa cum afirmă formula (4) în acest caz.

Cititorul va trebui să verifice următoarele exemple:

$k(x) = \sin\sqrt{x}$ $\qquad$ $k'(x) = (\cos\sqrt{x}) \cdot \frac{1}{2\sqrt{x}}$,

$k(x) = \sqrt{x} + \sqrt{x^5}$ $\qquad$ $k'(x) = (1 + 5x^2)\frac{1}{2\sqrt{x}}$,

$k(x) = \sin(x^2)$, $\qquad$ $k'(x) = \cos(x^2) \cdot 2x$,

$k(x) = \sin\frac{1}{x}$, $\qquad$ $k'(x) = -\cos\left(\frac{1}{x}\right)\frac{1}{x^2}$,

$k(x) = \sqrt{1 - x^2}$, $\qquad$ $k'(x) = \frac{-1}{2\sqrt{1 - x^2}} \cdot 2x = \frac{-x}{\sqrt{1 - x^2}}$.

Exerciţiu: Combinînd rezultatele de la p. 437 şi p. 458, arătaţi că funcţia

$$f(x) = \sqrt[m]{x^s} = x^{\frac{s}{m}}$$

are derivata

$$f'(x) = \frac{s}{m}x^{\frac{s}{m} - 1}.$$

Ar trebui să remarcăm că toate formulele referitoare la puterile lui x pot fi incluse într-una singură.

Dacă r este un număr raţional pozitiv sau negativ, atunci funcţia

$$f(x) = x^r$$

are derivata

$$f'(x) = rx^{r-1}.$$

Exerciții: 1) Efectuați derivările cerute în exercițiile de la p. 439, folosind regulile expuse în acest paragraf.

2) Derivați următoarele funcții:

$$x \sin x, \quad \frac{1}{1+x^2}\sin nx, \quad (x^3 - 3x^2 - x + 1)^3, \quad 1 + \sin^2 x, \quad x^2 \sin\frac{1}{x^2}, \quad \operatorname{arc\,sin}(\cos nx),$$

$$\operatorname{tg}\frac{1+x}{1-x}, \quad \operatorname{arc\,tg}\frac{1+x}{1-x}, \quad \sqrt[4]{1-x^2}, \quad \frac{1}{1+x^2}.$$

3) Găsiți derivatele de ordinul doi ale cîtorva din funcțiile precedente și ale funcțiilor

$$\frac{1-x}{1+x}, \quad \operatorname{arc\,tg} x, \quad \sin^2 x, \quad \operatorname{tg} x.$$

4) Derivați funcția $c_1(x - x_1)^2 + y_1^2 + c_2(x - x_2)^2 + y_2^2$ și demonstrați proprietățile de minim ale razei de lumină față de reflexie și de refracție, enunțate în capitolul VII, p. 348 și p. 398.

Reflecția sau refracția se fac în axa Ox, iar coordonatele extremităților drumului sînt respectiv x_1, y_1 și x_2, y_2.

(Observație: Funcția are un singur punct în care derivata se anulează; și deoarece este evident că avem un minim și nu un maxim, nu mai este necesar să studiem derivata de ordinul doi.)

Alte probleme referitoare a maxime și minime: 5) Găsiți extremele următoarelor funcții, schițați graficele lor, determinați intervalele de creștere, descreștere, convexitate și concavitate.

$$x^3 - 6x + 2, \quad x/(1 + x^2), \quad x^2/(1 + x^4), \quad \cos^2 x.$$

6) Studiați maximele și minimele funcției $x^3 + 3ax + 1$ și dependența lor de a.

7) Care din punctele hiperbolei $2y^2 - x^2 = 2$ este mai apropiat de punctul $x = 0$, $y = 3$?

8) Dintre toate dreptunghiurile de arie dată, găsiți-l pe acela care are cea mai scurtă diagonală.

9) Înscrieți dreptunghiul de arie maximă în elipsa $x^2/a^2 + y^2/b^2 = 1$.

10) Dintre toți cilindrii circulari de volum dat, găsiți-l pe acela de arie minimă.

§ 4. NOTAȚIA LUI LEIBNIZ ȘI „INFINITUL MIC”

Newton și Leibniz știau cum să obțină integrala și derivata ca limite. Dar bazele reale ale analizei au fost întunecate multă vreme datorită refuzului de a recunoaște dreptul exclusiv al noțiunii de limită ca sursă a noilor metode. Nici Newton și nici Leibniz nu s-au putut pronunța cu claritatea de astăzi, atît de simplă după ce noțiunea de limită a fost complet clarificată. Exemplul lor a dominat dezvoltarea matematicii mai bine de un secol, timp în care noțiunile erau învăluite cu cuvinte ca, de pildă, „cantități infinit mici”, „dife-

renţiale", „cîturi ultime" etc. Ezitarea cu care au fost părăsite aceste noţiuni în cele din urmă era adînc înrădăcinată în atitudinea filozofică a timpului şi chiar în natura intelectului uman. S-ar fi putut argumenta: Desigur, integrala şi derivata pot fi şi sînt calculate ca limite. Însă ce sînt, în definitiv aceste obiecte în sine, independent de modul particular în care ele sînt descrise prin procese de trecere la limită? Pare evident că noţiunile intuitive, ca de pildă aria sau panta unei curbe, au un înţeles absolut în sine, fără nici o nevoie de a fi explicate prin noţiuni auxiliare, ca de pildă, poligoane înscrise sau secante şi limitele lor. Într-adevăr, este natural din punct de vedere psihologic să căutăm definiţii adecvate pentru arie şi pantă, ca „lucruri în sine". Însă a renunţa la această dorinţă şi a vedea mai degrabă în procesele de trecere la limită singurele lor definiţii semnificative din punct de vedere ştiinţific este un fapt care concordă cu atitudinea matură care a deschis atît de des drumul spre progres. Însă în secolul al XVII-lea nu exista tradiţia intelectuală care ar fi putut permite acest radicalism filozofic.

Încercarea lui Leibniz de a „explica" derivata începe într-un mod foarte corect, prin formarea cîtului creşterilor unei funcţii $y = f(x)$,

$$\frac{\Delta y}{\Delta x} = \frac{f(x_1) - f(x)}{x_1 - x}.$$

Pentru limită, care este derivata pe care am notat-o $f'(x)$ (urmînd pe Lagrange, care a introdus această notaţie mai tîrziu), Leibniz scria

$$\frac{dy}{dx},$$

înlocuind simbolul diferenţei Δ prin „simbolul diferenţial" d. Dacă înţelegem faptul că acest simbol indică necesitatea efectuării procesului de trecere la limită $\Delta x \to 0$, şi deci $\Delta y \to 0$, nu întîmpinăm nici o dificultate sau mister. Înainte de a trece la limită, numitorul Δx din cîtul $\Delta y/\Delta x$ este simplificat sau transformat, astfel încît procesul de trecere la limită să poată fi efectuat cu uşurinţă. Aceasta este întotdeauna momentul crucial din procesul derivării. Dacă am fi încercat să trecem la limită înaintea acestei simplificări prealabile, am fi obţinut relaţia fără sens $\Delta y/\Delta x = 0/0$, care nu prezintă nici un interes. Misterul şi confuzia apar numai dacă urmăm pe Leibniz şi pe mulţi din continuatorii săi, spunînd următoarele:

„Δx nu tinde spre 0. În schimb, „ultima valoare" a lui Δx nu este zero, ci o „cantitate infinit mică", o „diferenţială" numită dx; în mod similar, Δy are o „ultimă" valoare infinit mică dy. Cîtul acestor diferenţiale infinit mici este un număr obişnuit $f'(x) = dy/dx$". În consecinţă, Leibniz numea derivata *cît diferenţial*. Aceste cantităţi infinit mici erau considerate ca noi feluri de numere, nenule, dar mai mici decît orice număr pozitiv real. Numai aceia care aveau un adevărat simţ matematic puteau sesiza această noţiune, iar ana-

liza matematică era considerată a fi deosebit de dificilă, deoarece nu oricine are sau își poate dezvolta acest simț. În același mod, integrala era considerată ca sumă a unei infinități de „cantități infinit mici" $f(x)\mathrm{d}x$. O astfel de sumă, după cum încerca să-și închipuie lumea, este integrala sau aria, în timp ce calculul valorii ei ca ***limită a unei sume finite de numere obișnuite*** $f(x_j)\Delta x$ era privită ca fiind ceva auxiliar. Astăzi renunțăm, pur și simplu, la dorința unei explicații „directe" și definim integrala ca limită a unei sume finite. În acest mod, dificultățile sînt risipite și tot ceea ce este de preț în analiză se așază pe o bază solidă.

În ciuda acestei dezvoltări ulterioare, notația lui Leibniz, $\mathrm{d}y/\mathrm{d}x$ pentru derivata $f'(x)$ și $\int f(x)\mathrm{d}x$ pentru integrală au fost păstrate și s-au dovedit a fi deosebit de utile. Nu este nici o supărare dacă considerăm simbolurile d doar ca simboluri pentru trecerea la limită. Notația lui Leibniz are avantajul că limitele cîturilor și ale sumelor pot fi tratate într-o oarecare măsură, „ca și cum" ele ar fi cîturi sau sume adevărate. Puterea sugestivă a acestui simbolism a îndemnat întotdeauna lumea să atribuie acestor simboluri un înțeles cu totul nematematic. Dacă rezistăm acestei tentații, atunci notația lui Leibniz este cel puțin o prescurtare excelentă a notației explicite mai greoaie, a procesului de trecere la limită; de fapt, ea este aproape indispensabilă în părțile mai avansate ale teoriei.

De exemplu, regula (d) de la p. 448 pentru derivarea funcției inverse $x = g(y)$ a funcției $y = f(x)$ era $g'(y)f'(x) = 1$. Cu notația lui Leibniz ea se scrie

$$\frac{\mathrm{d}x}{\mathrm{d}y} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = 1$$

„ca și cum" „diferențialele" ar putea fi simplificate, ca într-o fracție obișnuită. În mod analog, regula (e) de la p. 449 pentru derivarea unei funcții compuse $z = k(x)$, unde

$$z = g(y), \quad y = f(x),$$

se scrie acum

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}y}\,\frac{\mathrm{d}y}{\mathrm{d}x}.$$

Notația lui Leibniz are avantajul sublinierii mai degrabă a cantităților x, y, z decît a legăturii funcționale explicite dintre ele. Această legătură funcțională exprimă un *procedeu*, o *operație* cu ajutorul căreia dintr-o cantitate x se obține alta y; de pildă, funcția $y = f(x) = x^2$, determină o cantitate y, egală cu pătratul cantității x. Operația (ridicarea la pătrat) este obiectul atenției matematicianului. Fizicienii și inginerii sînt interesați, în primul rînd, în cunoașterea cantităților. De aceea, accentul pe care notația lui Leibniz îl

pune asupra cantităților are o atracție deosebită pentru cei care se ocupă cu matematica aplicată.

Mai putem face o observație. În timp ce „diferențialele" considerate drept cantități mici sînt acum căzute în dizgrație în mod definitiv, același cuvînt „diferențială" s-a strecurat prin ușa din spate, pentru a desemna de data aceasta o noțiune perfect legitimă și folositoare. Acum ea înseamnă o diferență Δx, cînd Δx este mică în comparație cu celelalte cantități care intervin[2]. Nu putem discuta mai amănunțit importanța acestei noțiuni pentru calculele aproximative. Nu vom discuta nici alte noțiuni matematice legitime, pentru care cuvîntul „diferențial" a fost adoptat, dintre care unele s-au dovedit a fi deosebit de folositoare în analiză și în aplicațiile ei în geometrie.

§ 5. TEOREMA FUNDAMENTALĂ A ANALIZEI

1. Teorema fundamentală

Noțiunea de integrare și, într-o oarecare măsură, aceea de derivare au fost dezvoltate destul de bine înaintea lucrărilor lui Newton și Leibniz. Pentru a porni extraordinara evoluție a noii analize matematice, mai era necesară doar o descoperire foarte simplă. Cele două procese de trecere la limită care intervin în derivarea și integrarea unei funcții, aparent nelegate între ele, sînt în strînsă legătură. Într-adevăr, ele sînt operații inverse, tot astfel cum sînt operațiile de adunare și scădere, sau înmulțire și împărțire. Nu există un calcul diferențial și unul integral separat, ci un singur calcul[3].

Marele merit al lui Leibniz și Newton constă în faptul că ei au recunoscut clar și au folosit pentru prima dată această *teoremă fundamentală a analizei*. Desigur descoperirea lor se află pe drumul drept al dezvoltării științifice și a fost foarte natural ca mai mulți oameni să ajungă la o înțelegere clară a situației, în mod independent și aproape simultan. Pentru a formula teorema fundamentală, să considerăm integrala unei funcții $y = f(x)$ de la limita inferioară fixată a la limita superioară variabilă x. Pentru a evita confuzia dintre limita superioară de integrare x și variabila x care apare în simbolul $f(x)$, să scriem această integrală sub forma (cf. p. 429)

$$F(x) = \int_a^x f(u)\, du, \tag{1}$$

[2] Semnificația modernă a noțiunii de „diferențială" nu este chiar aceasta. Cititorul poate consulta în acest sens acad. M. NICOLESCU, N. DINCULEANU, S. MARCUS, *Analiza matematică*, vol. I, Ed. didactică și pedagogică, București, 1966 — *N.T.*

[3] În limba engleză se folosește termenul *calculus*, care înseamnă calcul, pentru a desemna disciplina matematică cunoscută în limba română sub termenul „calcul diferențial și integral". Termenul „analiză matematică" denotă o disciplină matematică mai cuprinzătoare decît prima, care se referă mai degrabă la aspectul algoritmic al analizei matematice. — *N.T.*

indicînd faptul că dorim să studiem integrala ca o funcţie $F(x)$ de limita superioară x (fig. 274). Această funcţie $F(x)$ este aria aflată sub curba $y = f(u)$, de la punctul $u = a$ la punctul $u = x$. Uneori integrala $F(x)$ cu limita superioară variabilă se numeşte integrală „nedefinită".

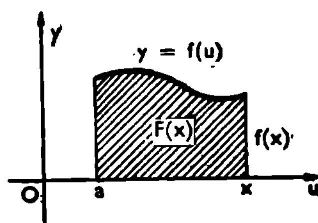


Fig. 274. Integrala ca funcţie de limita superioară

Acum teorema fundamentală a calculului diferenţial şi integral este următoarea:

Derivata integralei nedefinite (1) ca funcţie x *este egală cu valoarea funcţiei* f(u) *în punctul* x:

$$F'(x) = f(x).$$

Cu alte cuvinte, procesul de integrare care duce de la funcţia f(x) la F(x) *este inversat prin procesul de derivare aplicat lui* F(x).

Pe o bază intuitivă, demonstraţia este foarte simplă. Ea depinde de interpretarea integralei $F(x)$ ca arie, dar ne-am îndepărta mult de scopul propus, dacă am încerca să reprezentăm pe $F(x)$ printr-un grafic şi derivata $F'(x)$ prin panta lui. În locul interpretării geometrice iniţiale a derivatei reţinem explicaţia geometrică a integralei $F(x)$, dar procedăm într-un mod analitic pentru derivarea lui $F(x)$. Diferenţa

$$F(x_1) - F(x)$$

este pur şi simplu aria cuprinsă între x şi x_1 din fig. 275, şi vedem că această arie se află între valorile $(x_1 - x)m$ şi $(x_1 - x)M$,

$$(x_1 - x)m \leqslant F(x_1) - F(x) \leqslant (x_1 - x)M,$$

unde M şi m sînt respectiv cea mai mare şi cea mai mică valoare a lui $f(u)$ în intervalul cuprins între x şi x_1. Într-adevăr, aceste două produse sînt ariile dreptunghiurilor care includ, respectiv sînt incluse, în trapezul curbiliniu. De aceea

$$m \leqslant \frac{F(x_1) - F(x)}{x_1 - x} \leqslant M.$$

Vom presupune că funcţia $f(u)$ este continuă, astfel încît dacă x_1 tinde către x, atunci M şi m tind ambele către $f(x)$

Deci avem

$$F'(x) = \lim \frac{F(x_1) - F(x)}{x_1 - x} = f(x), \tag{2}$$

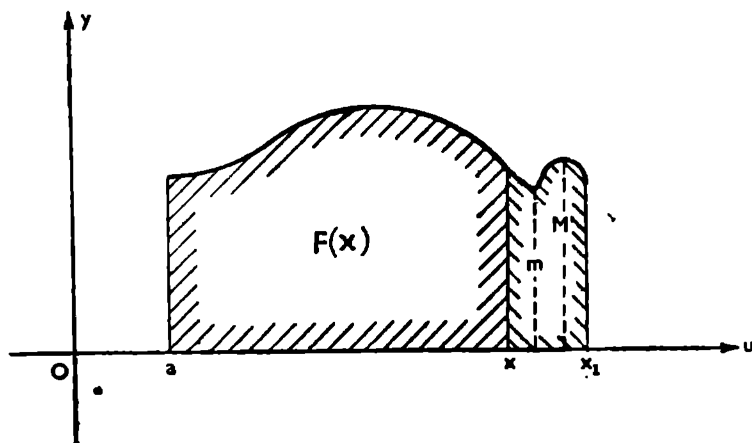


Fig. 275. Demonstrarea teoremei fundamentale

după cum am afirmat. Intuitiv, această formulă exprimă faptul că viteza de variaţie a ariei de sub curba $y = f(x)$, atunci cînd x creşte, este egală cu înălţimea curbei în punctul x.

În anumite manuale, momentul esenţial al teoremei fundamentale este ascuns de o nomenclatură prost aleasă. Mulţi autori introduc mai întîi derivata şi apoi definesc ,,integrala nedefinită'' pur şi simplu ca inversă a derivatei, spunînd că $G(x)$ este o integrală nedefinită a lui $f(x)$, dacă

$$G'(x) = f(x).$$

În acest mod, această metodă combină imediat derivarea cu cuvîntul ,,integrală''. Abia mai tîrziu se introduce noţiunea de ,,integrală definită'' ca arie sau ca limită a unei sume, fără a sublinia faptul că acum cuvîntul ,,integrală'' are un înţeles cu totul deosebit. În acest mod, faptul cel mai important al teoriei este introdus prin uşa din spate şi studentul este împiedicat în mod serios în eforturile sale îndreptate spre înţelegerea clară a lucrurilor. Preferăm să numim funcţiile $G(x)$, pentru care $G'(x) = f(x)$, nu prin termenul ,,integrale nedefinite'', ci *funcţii primitive* ale lui $f(x)$. Teorema fundamentală se enunţă atunci în modul următor:

F(x), *integrala lui* f(u) *cu limita inferioară fixată și cu limita superioară variabilă* x, *este o funcție primitivă a funcției* f(x).

Spunem „o funcție primitivă" și nu „funcția primitivă", deoarece se observă imediat că dacă $G(x)$ este o funcție primitivă a lui $f(x)$, atunci:

$H(x) = G(x) + c$ (c fiind o constantă arbitrară) este și ea o primitivă, deoarece $H'(x) = G'(x)$. Reciproca este și ea adevărată. Două funcții primitive $G(x)$ și $H(x)$ *diferă numai printr-o constantă, pentru că diferența* $U(x) = G(x) - H(x)$ are derivata $U'(x) = G'(x) - H'(x) = f(x) - f(x) = 0$ și este de aceea constantă, deoarece o funcție, al cărei grafic este orizontal peste tot, trebuie să fie constantă.

Acest fapt duce la o regulă foarte importantă pentru găsirea valorii unei integrale de la a la b, cu condiția să cunoaștem o funcție primitivă $G(x)$ a lui $f(x)$. În conformitate cu teorema fundamentală funcția

$$F(x) = \int_a^x f(u)du,$$

este și ea o funcție primitivă a funcției $f(x)$. Deci $F(x) = G(x) + c$ unde c este o constantă. Putem determina constanta c dacă ne reamintim că $F(a) = \int_a^a f(u)du = 0$. Aceasta ne dă $0 = G(a) + c$, astfel încît $c = -G(a)$. Atunci integrala definită între limitele a și x va fi $F(x) = \int_a^x f(u)du = G(x) - G(a)$, sau dacă scriem b în loc de x

$$\int_a^b f(u)du = G(b) - G(a), \tag{3}$$

independent de primitiva particulară $G(x)$, pe care am ales-o. Cu alte cuvinte:

Pentru a evalua integrala definită $\int_a^b f(x)dx$ *trebuie doar să găsim o funcție* G(x) *astfel încît* G′(x) = f(x) *și să formăm apoi diferența* G(b) — G(a).

2. Primele aplicații. Integrarea funcțiilor x^r, $\cos x$, $\sin x$, $\operatorname{arc\,tg} x$

Nu putem da aici o imagine potrivită a importanței teoremei fundamentale, însă exemplificările care urmează constituie o indicație destul de bună. În problemele efective întîlnite în mecanică, fizică sau în matematica pură, se

cere adesea valoarea unei integrale definite. Încercarea directă de a găsi integrala ca limită a unei sume poate fi dificilă. Pe de altă parte, după cum am văzut în § 3, este destul de ușor să efectuăm orice derivare și să acumulăm o mare cantitate de informație în acest domeniu. Orice formulă de derivare $G'(x) = f(x)$ poate fi citită în sens invers, și atunci ne furnizează o funcție primitivă $G(x)$ a funcției $f(x)$. Cu ajutorul formulei (3), acest lucru poate fi folosit pentru calculul integralei funcției $f(x)$, între două limite oarecare.

De exemplu, dacă vrem să găsim integralele funcțiilor x^2, sau x^3 sau x^n, putem proceda mult mai simplu decît în § 1. Știm din formula de derivare pentru x^n că derivata lui x^n este nx^{n-1}, astfel încît derivata funcției:

$$G(x) = \frac{x^{n+1}}{n+1} \quad (n \neq -1).$$

este

$$G'(x) = \frac{n+1}{n+1} x^n = x^n.$$

De aceea $x^{n+1}/(n+1)$ este o funcție primitivă a funcției $f(x) = x^n$ și deci avem imediat

$$\int_a^b x^n \, \mathrm{d}x = G(b) - G(a) = \frac{b^{n+1} - a^{n+1}}{n+1}.$$

Acest procedeu este mult mai simplu decît acela prin care integrala se obține ca limită a unei sume.

Mai general, am găsit în § 3, că pentru orice număr rațional s, pozitiv sau negativ, funcția x^s are derivata sx^{s-1} și de aceea pentru $s = r + 1$, funcția

$$G(x) = \frac{1}{r+1} x^{r+1}$$

are derivata $f(x) = G'(x) = x^r$. (Presupunem că $r \neq -1$, adică $s \neq 0$.) Deci $x^{r+1}/(r+1)$ este o funcție primitivă a funcției x^r și avem (pentru a, b pozitive și $r \neq -1$)

$$\int_a^b x^r \mathrm{d}x = \frac{1}{r+1} (b^{r+1} - a^{r+1}). \tag{4}$$

În (4) am presupus că în intervalul de integrare integrandul x^r este definit și continuu, ceea ce exclude valoarea $x = 0$ dacă $r < 0$. De aceea, facem ipoteza că în acest caz a și b sînt pozitive.

Pentru $G(x) = -\cos x$ avem $G'(x) = \sin x$, deci

$$\int_0^a \sin x \, dx = -(\cos a - \cos 0) = 1 - \cos a.$$

În mod similar, deoarece pentru $G(x) = \sin x$ avem $G'(x) = \cos x$ rezultă că

$$\int_0^a \cos x \, dx = \sin a - \sin 0 = \sin a.$$

Se obţine un rezultat deosebit de interesant din formula de derivare a funcţiei arc tg, $D \operatorname{arc\,tg} x = 1/(1 + x^2)$. Rezultă că funcţia arc tg x este o funcţie primitivă a funcţiei $1/(1 + x^2)$, şi din formula (3) obţinem rezultatul

$$\operatorname{arc\,tg} b - \operatorname{arc\,tg} 0 = \int_0^b \frac{1}{1+x^2} \, dx.$$

Dar arc tg 0 = 0, pentru că valorii 0 a tangentei i se asociază valoarea 0 a unghiului. Deci găsim

(5)
$$\operatorname{arc\,tg} b = \int_0^b \frac{1}{1 + x^2} \, dx.$$

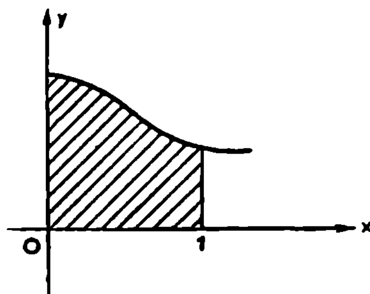


Fig. 276. $\pi/4$ ca arie mărginită de curba $y = 1/(1 + x^2)$ între 0 şi 1

Dacă în particular $b = 1$, atunci arc tg b va fi egal cu $\pi/4$, deoarece valorii 1 a tangentei îi corespunde un unghi de 45°, care măsurat în radiani este egal cu $\pi/4$. Obţinem deci formula remarcabilă

(6)
$$\pi/4 = \int_0^1 \frac{1}{1 + x^2} \, dx.$$

Aceasta arată că aria aflată sub graficul funcţiei $y = 1/(1 + x^2)$ de la $x = 0$ la $x = 1$ este egală cu un sfert din aria cercului de rază 1.

3. Formula lui Leibniz pentru numărul π

Ultimul rezultat ne conduce la una dintre cele mai frumoase descoperiri matematice făcute în secolul al XVII-lea — seria alternată a lui Leibniz pentru numărul π:

(7) $$\frac{\pi}{4} = \frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \dots .$$

Prin simbolul $+ \dots$ înțelegem că șirul de „sume parțiale" finite, format prin secționarea expresiei din membrul drept după primii n termeni, converge spre limita $\pi/4$, cînd n crește.

Pentru a demonstra această formulă celebră, este suficient să ne reamintim formula sumei unei progresii geometrice finite

$$\frac{1-q^n}{1-q} = 1 + q + q^2 + \dots + q^{n-1} \text{ sau } \frac{1}{1-q} = 1 + q + q^2 + \dots + q^{n-1} + \frac{q^n}{1-q}$$

În această identitate algebrică facem substituția $q = -x^2$ și obținem

(8) $$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots + (-1)^{n-1} x^{2n-2} + R_n,$$

în care „restul" R_n este

$$R_n = (-1)^n \frac{x^{2n}}{1+x^2}.$$

Egalitatea (8) poate fi acum integrată între limitele 0 și 1. Din regula (a) din § 3 rezultă că trebuie să calculăm suma integralelor termenilor din membrul drept. Deoarece în baza formulei (4) avem $\int_a^b x^m dx = (b^{m+1} - a^{m+1})/(m+1)$,

găsim că

$$\int_0^1 x^m dx = 1/(m+1) \text{ și de aceea:}$$

(9) $$\int_0^1 \frac{dx}{1+x^2} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots + (-1)^{n-1} \frac{1}{2n-1} + T_n,$$

unde

$$T_n = (-1)^n \int_0^1 \frac{x^{2n}}{1+x^2}\,dx.$$

Conform formulei (5), membrul stîng din (9) este egal cu $\pi/4$. Diferența dintre $\pi/4$ și suma parțială

$$S_n = 1 - \frac{1}{3} + \frac{1}{5} + \ldots + \frac{(-1)^{n-1}}{2n-1}$$

este $\pi/4 - S_n = T_n$. Rămîne de arătat că T_n tinde spre 0, cînd n crește. Dar

$$\frac{x^{2n}}{1+x^2} \leqslant x^{2n} \qquad \text{pentru } 0 \leqslant x \leqslant 1.$$

Reamintindu-ne formula (13) din § 1, care afirmă că $\int_a^b f(x)dx \leqslant \int_a^b g(x)dx$, dacă $f(x) \leqslant g(x)$ și $a < b$, vedem că

$$|T_n| = \int_0^1 \frac{x^{2n}}{1+x^2}\,dx \leqslant \int_0^1 x^{2n}\,dx;$$

deoarece membrul drept este egal cu $1/(2n+1)$, așa cum am văzut mai înainte (formula (4)), găsim $|T_n| < 1/(2n+1)$. Deci

$$\left|\frac{\pi}{4} - S_n\right| < \frac{1}{2n+1}.$$

Însă aceasta arată că S_n tinde spre $\pi/4$, cînd n crește, deoarece $1/(2n+1)$ tinde spre zero. Astfel, formula lui Leibniz este demonstrată.

§ 6. FUNCȚIA EXPONENȚIALĂ ȘI LOGARITMUL

Noțiunile fundamentale ale analizei permit o teorie mai adecvată a logaritmului și a funcției exponențiale decît aceea predată în școală prin procedee „elementare". De obicei, se începe cu puterile întregi a^n ale unui număr pozitiv a și după aceea se definește $a^{1/m} = \sqrt[m]{a}$, obținînd după aceea valorile lui a^r pentru orice număr rațional $r = n/m$. Apoi se definește valoarea lui a^x pentru orice x irațional, astfel încît a^x să fie funcție continuă de x, moment delicat care este evitat în instrucția elementară.

În sfîrșit, logaritmul lui y, în baza a,

$$x = \log_a y,$$

este definit ca funcție inversă a funcției $y = a^x$.

În teoria acestor funcții, bazată pe calcul diferențial și integral și care se va face în cele ce urmează, se inversează ordinea în care ele sînt considerate. Începem cu logaritmul și apoi obținem funcția exponențială.

1. Definiția și proprietățile logaritmului. Numărul e al lui Euler

Definim logaritmul, sau mai precis „logaritmul natural" $F(x) = \log x$ (legătura lui cu logaritmul în baza 10 va fi arătată în secțiunea 2), ca fiind aria mărginită de curba $y = 1/u$, de la $u = 1$ la $u = x$, sau, ceea ce este același lucru, prin integrala

$$F(x) = \log x = \int_1^x \frac{1}{u}\, du \tag{1}$$

(fig. 5). Variabila x poate fi orice număr pozitiv. Zero este exclus, deoarece integrandul $1/u$ devine infinit, cînd u tinde către zero.

Este foarte natural să studiem funcția $F(x)$, pentru că știm că primitiva oricărei puteri x^n este funcția $x^{n+1}/(n+1)$ de același tip, cu excepția cazului în care $n = -1$. În ultimul caz, numitorul $n + 1$ s-ar anula și formula (4) de la p. 458 nu ar avea sens. Prin urmare, ne putem aștepta ca integrarea funcției $1/x$ sau $1/u$ să ne ducă la un nou tip interesant de funcție.

Cu toate că definim funcția $\log x$ prin formula (1), nu „cunoaștem" funcția decît după ce am dedus proprietățile ei și am găsit metode pentru calculul ei numeric. Este un exemplu caracteristic pentru metodele moderne să începem cu noțiuni generale, ca de pildă aria și integrala, să dăm pe această bază definiții, ca de pildă cea din formula (1), și să deducem apoi proprietățile obiectelor definite, obținînd abia la sfîrșit expresii explicite pentru calculul numeric.

Prima proprietate importantă a funcției $\log x$ este o consecință imediată a teoremei fundamentale din § 5. Această teoremă ne conduce la egalitatea

$$F'(x) = 1/x. \tag{2}$$

Din (2) rezultă că derivata este mereu pozitivă, ceea ce confirmă faptul evident că funcția $\log x$ este o funcție strict crescătoare, atunci cînd ne deplasăm în sensul valorilor crescătoare ale lui x.

Proprietatea principală a logaritmului este exprimată prin formula

$$\log a + \log b = \log (ab). \tag{3}$$

Importanța acestei formule în aplicațiile practice ale logaritmului la calcule numerice este binecunoscută. Formula (3) ar putea fi obținută și intuitiv, observînd ariile care definesc cele trei cantități $\log a$, $\log b$ și $\log (ab)$. Preferăm însă să o deducem printr-un raționament caracteristic analizei: o dată cu funcția $F(x) = \log x$, să considerăm și funcția:

$$k(x) = \log (ax) = \log w = F(w),$$

punînd $w = f(x) = ax$, unde a este o constantă pozitivă. Putem deriva cu ușurință funcția $k(x)$ folosind regula (e) din § 3; $k'(x) = F'(w)f'(x)$. Din (2), ținînd seama că $f'(x) = a$, deducem

$$k'(x) = a/w = a/ax = 1/x.$$

Prin urmare, funcția $k(x)$ are aceeași derivată ca și $F(x)$, deci, conform celor arătate la p. 456 avem

$$\log (ax) = k(x) = F(x) + c,$$

unde c este o constantă, care nu depinde de valoarea particulară a lui x. Constanta c se determină prin procedeul simplu al înlocuirii lui x cu numărul 1. Din definiția (1) știm că

$$F(1) = \log 1 = 0,$$

pentru că integrala are limitele superioară și inferioară egale, cînd $x = 1$. Deci obținem

$$k(1) = \log (a \cdot 1) = \log a = \log 1 + c = c,$$

ceea ce dă $c = \log a$ și de aceea pentru orice x avem formula:

(3a) $$\log (ax) = \log a + \log x.$$

Punînd $x = b$, obținem formula dorită (3).

În particular (pentru $a = x$), obținem pe rînd

(4) $$\log(x^2) = 2 \log x,$$

$$\log(x^3) = 3 \log x,$$

$$\cdots\cdots$$

$$\log(x^n) = n \log x.$$

Egalitatea (4) ne arată că pentru valori crescătoare ale lui x valorile funcției $\log x$ tind spre infinit, pentru că logaritmul este o funcție strict crescătoare și avem, de pildă,

$$\log(2^n) = n \log 2,$$

care tinde la infinit o dată cu n. Mai mult avem

$$0 = \log 1 = \log\left(x \cdot \frac{1}{x}\right) = \log x + \log \frac{1}{x},$$

astfel încît

(5) $$\log \frac{1}{x} = -\log x.$$

În sfîrşit,

(6) $$\log x^r = r \log x$$

pentru orice număr raţional $r = m/n$, pentru că, punînd $x^r = u$, avem:

$$n \log u = \log u^n = \log x^{\frac{m}{n} \cdot n} = \log x^m = m \log x,$$

astfel încît

$$\log x^{\frac{m}{n}} = \frac{m}{n} \log x.$$

Deoarece $\log x$ este o funcţie continuă şi crescătoare de x, care se anulează pentru $x = 1$ şi tinde spre infinit cînd x creşte, trebuie să existe un număr mai mare decît 1, astfel încît pentru această valoare să avem: $\log x = 1$. Urmînd

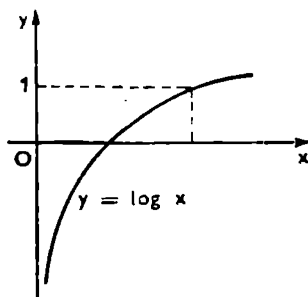


Fig. 277.

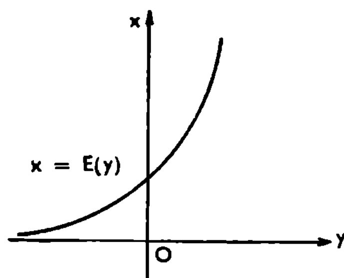


Fig. 278.

calea indicată de Euler, să notăm acest număr cu e. (Echivalenţa cu definiţia de la p. 315 va fi arătată mai tîrziu.) Astfel e este definit prin egalitatea

(7) $$\log e = 1.$$

Am introdus numărul e printr-o proprietate intrinsecă, care asigură existenţa lui. Acum vom continua această analiză, obţinînd ca o consecinţă formule explicite, care dau aproximaţii oricît de exacte pentru valoarea numerică a lui e.

2. Funcţia exponenţială

Rezumînd rezultatele precedente, vedem că funcţia $F(x) = \log x$ se anulează pentru $x = 1$, creşte monoton tinzînd spre infinit, dar cu panta descrescătoare $1/x$, iar pentru valori pozitive ale lui x, mai mici decît 1, este dată de opusul lui $\log 1/x$, astfel încît $\log x$ tinde spre $-\infty$ atunci cînd $x \to 0$.

Datorită caracterului monoton al funcţiei $y = \log x$, putem considera funcţia inversă

$$x = E(y),$$

al cărei grafic (fig. 278) se obţine în modul obişnuit din graficul funcţiei $y = \log x$ (fig. 277) şi care este definită pentru toate valorile y cuprinse între $-\infty$ şi $+\infty$. Cînd y tinde spre $-\infty$, valoarea $E(y)$ tinde spre zero, iar cînd y tinde spre $+\infty$, $E(y)$ tinde spre $+\infty$.

Funcţia E are următoarea proprietate fundamentală:

$$E(a) \cdot E(b) = E(a + b) \tag{8}$$

pentru orice pereche de valori a şi b. Această lege este doar o altă formă a formulei (3) pentru logaritm. Într-adevăr, dacă punem

$$E(b) = x, \quad E(a) = z \quad (\text{adică } b = \log x,\ a = \log z),$$

avem

$$\log xz = \log x + \log z = b + a,$$

şi de aceea

$$E(b + a) = xz = E(a) \cdot E(b),$$

ceea ce trebuia demonstrat.

Deoarece prin definiţie $\log e = 1$, avem

$$E(1) = e,$$

iar din (8) rezultă că $e^2 = E(1)\,E(1) = E(2)$ etc. În general,

$$E(n) = e^n \text{ pentru orice număr întreg } n.$$

De asemenea, $E(1/n) = e^{1/n}$, astfel încît $E(p/q) = E(1/q) \dots E(1/q) = (e^{1/q})^p$; deci punînd $p/q = r$, avem

$$E(r) = e^r$$

pentru orice număr raţional r. De aceea este potrivit să *definim* operaţia de ridicare a numărului e la o putere iraţională, punînd

$$e^y = E(y)$$

pentru orice număr real y, deoarece funcţia E este continuă pentru toate valorile lui y şi coincide cu valorile lui e^y pentru y raţional. Acum putem exprima legea fundamentală (8) a funcţiei E, sau a *funcţiei exponenţiale* cum mai este numită, prin egalitatea

$$e^a e^b = e^{a+b}, \tag{9}$$

care este acum stabilită pentru orice numere a şi b, raţionale sau iraţionale.

În discuţiile de mai sus ne-am referit la funcţiile logaritmică şi exponenţială cu „baza" e, „baza naturală" a logaritmului. Trecerea de la baza e la orice alt număr pozitiv se face cu uşurinţă. Începem prin a considera logaritmul (natural)

$$\alpha = \log a,$$

astfel încît

$$a = e^\alpha = e^{\log a}.$$

Acum să definim a^x prin expresia compusă

$$z = a^x = e^{\alpha x} = e^{x \log a}. \tag{10}$$

De exemplu,

$$10^x = e^{x \log 10}.$$

Numim inversa funcţiei a^x *logaritmul cu baza a* şi vedem imediat că *logaritmul natural* al lui z este egal cu produsul dintre x şi α; cu alte cuvinte, logaritmul cu baza a al unui număr z se obţine împărţind logaritmul natural al lui z prin logaritmul natural fixat al lui a*. Pentru $a = 10$, aceasta este

$$\log 10 = 2{,}303.$$

(cu patru cifre).

3. Formule pentru derivarea funcţiilor e^x, a^x, x^s

Deoarece am definit funcţia exponenţială $E(y)$ ca fiind inversa funcţiei $y = \log x$, din regula de derivare a funcţiilor inverse (§ 3) rezultă că

$$E'(y) = \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{\frac{1}{x}} = x = E(y),$$

adică

$$E'(y) = E(y). \tag{11}$$

Funcţia exponenţială naturală este deci egală cu propria ei derivată.

---

* În aceste consideraţii facem presupunerea că $a \neq 1$. — *N.T.*

Aceasta este într-adevăr sursa tuturor proprietăților funcției exponențiale și motivul fundamental al importanței ei pentru aplicații, așa cum se va vedea în secțiunile următoare. Folosind notația introdusă în §2, putem scrie formula (11) sub forma

$$\frac{d}{dx} e^x = e^x. \tag{11 a}$$

Mai general, derivînd funcția compusă

$$f(x) = e^{\alpha x},$$

cu ajutorul regulei din §3, obținem

$$f'(x) = \alpha e^{\alpha x} = \alpha f(x).$$

Deci, pentru $\alpha = \log a$, găsim că funcția

$$f(x) = a^x$$

are derivata

$$f'(x) = a^x \log a.$$

Acum putem defini funcția

$$f(x) = x^s$$

pentru orice exponent real s și orice variabilă pozitivă x, punînd

$$x^s = e^{s \log x}.$$

Aplicînd din nou regula de derivare a funcțiilor compuse, $f(x) = e^{sz}$, $z = \log x$, găsim că $f'(x) = se^{sz} \cdot 1/x = sx^s \cdot 1/x$ și de aceea

$$f'(x) = sx^{s-1}$$

conform rezultatului precedent, obținut pentru valori raționale ale lui s.

4. Expresii explicite petru e, e^x și $\log x$ cu ajutorul limitelor

Pentru a găsi formule explicite pentru aceste funcții, vom folosi formulele de derivare ale funcției exponențiale și logaritmice. Deoarece derivata funcției $\log x$ este egală cu $1/x$, din definiția derivatei obținem relația

$$\frac{1}{x} = \lim \frac{\log x_1 - \log x}{x_1 - x}, \text{ cînd } x_1 \to x.$$

Dacă punem $x_1 = x + h$ și facem ca h să tindă spre zero, luînd valorile

$$h = 1/2,\ 1/3,\ 1/4,\ \ldots,\ 1/n,\ \ldots,$$

atunci, aplicînd regulile logaritmului, găsim

$$\frac{\log\left(x+\frac{1}{n}\right)-\log x}{1/n}=n\log\frac{x+\frac{1}{n}}{x}=\log\left[\left(1+\frac{1}{nx}\right)^n\right]\to\frac{1}{x}$$

Scriind $z = 1/x$ şi folosind din nou legile logaritmului, obţinem

$$z=\lim\log\left[\left(1+\frac{z}{n}\right)^n\right]\text{ pentru } n\to\infty.$$

Utilizînd funcţia exponenţială, găsim

$$e^z=\lim\left(1+\frac{z}{n}\right)^n \text{ pentru } n\to\infty. \tag{12}$$

Am obţinut, prin urmare, celebra formulă care defineşte funcţia exponenţială printr-o limită. În particular, pentru $z = 1$, găsim

$$e=\lim\,(1+1/n)^n, \tag{13}$$

iar pentru $z = -1$,

$$\frac{1}{e}=\lim\,(1-1/n)^n. \tag{13a}$$

Aceste expresii conduc îndată la dezvoltări sub forma unor serii infinite. Din teorema binomului găsim că

$$\left(1+\frac{x}{n}\right)^n=1+n\frac{x}{n}+\frac{n(n+1)}{2\,!}\cdot\frac{x^2}{n^2}+\frac{n(n-1)(n-2)}{3\,!}\cdot\frac{x^3}{n^3}+ \\ +\ldots+\frac{x^n}{n^n}$$

sau

$$\left(1+\frac{x}{n}\right)^n=1+\frac{x}{1\,!}+\frac{x^2}{2\,!}\left(1-\frac{1}{n}\right)+\frac{x^3}{3\,!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+ \\ +\ldots+\frac{x^n}{n\,!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{n-2}{n}\right)\left(1-\frac{n-1}{n}\right).$$

Este plauzibil şi nu este greu de dat o justificare completă (omitem aici amănuntele) a faptului că putem efectua trecerea la limită cînd $n\to\infty$, înlocuind pe $1/n$ prin 0 în fiecare termen. Aceasta ne dă celebra serie infinită pentru e^x

$$e^x=1+\frac{x}{1\,!}+\frac{x^2}{2\,!}+\frac{x^3}{3\,!}+\ldots, \tag{14}$$

şi, în particular, seria pentru e

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots,$$

ceea ce stabileşte identitatea lui e cu numărul definit la p. 315. Pentru $x = -1$, obţinem seria

$$\frac{1}{e} = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \dots,$$

care dă o aproximare numerică excelentă cu foarte puţini termeni, eroarea totală efectuată prin secţionarea seriei după cel de al n-lea termen, fiind mai mică decît primul termen neglijat.

Folosind formula de derivare a funcţiei exponenţiale, putem obţine o expresie interesantă pentru logaritm. Avem

$$\lim \frac{e^h - 1}{h} = \lim \frac{e^h - e^0}{h} = 1,$$

cînd h tinde spre zero, deoarece această limită este derivata funcţiei e^y pentru $y = 0$ şi ea este egală cu $e^0 = 1$. În această formulă punem în locul lui h valorile z/n, unde z este un număr arbitrar iar n parcurge şirul întregilor pozitivi. Aceasta dă

$$n \frac{e^{z/n} - 1}{z} \to 1$$

sau

$$n(\sqrt[n]{e^z} - 1) \to z,$$

cînd n tinde spre infinit. Scriind $z = \log x$ sau $e^z = x$, obţinem în cele din urmă,

$$\log x = \lim n(\sqrt[n]{x} - 1), \text{ cînd } n \to \infty. \tag{15}$$

Deoarece $\sqrt[n]{x} \to 1$ cînd $n \to \infty$ (cf. p. 341), această formulă ne dă logaritmul ca limită a unui produs în care unul dintre factori tinde spre zero, iar celălalt spre infinit.

Diferite exemple şi exerciţii: Incluzînd funcţia exponenţială şi cea logaritmică, stăpînim acum o clasă largă de funcţii şi putem face multe aplicaţii:

Derivaţi funcţiile: 1) $x(\log x - 1)$; 2) $\log(\log x)$; 3) $\log(x + \sqrt{1 + x^2})$; 4) $\log(x + \sqrt{1 - x^2})$; 5) e^{-x^2}; 6) e^{e^x} (O funcţie compusă e^z, cu $z = e^x$.); 7) x^x (Indicaţie: $x^x = e^{x \log x}$); 8) $\log \operatorname{tg} x$; 9) $\log \sin x$; $\log \cos x$; 10) $x/\log x$.

Găsiţi maximele şi minimele funcţiilor 11) xe^{-x}; 12) x^2e^{-x}; 13) xe^{-ax}.

*14) Găsiţi locul geometric al punctului maxim al curbei $y = xe^{-ax}$ cînd a variază.

*15) Arătaţi că toate derivatele succesive ale lui e^{-x^2} au forma produsului dintre e^{-x^2} şi un polinom în x.

*16) Arătaţi că derivata de ordinul n a funcţiei e^{-1/x^2} este de forma produsului dintre funcţia $e^{-1/x^2} \cdot 1/x^{3n}$ cu un polinom de gradul $2n - 2$.

*17) *Derivarea logaritmică.* Folosind proprietatea fundamentală a logaritmului, derivarea produselor poate fi efectuată uneori într-un mod mai simplu. Pentru un produs de forma

$$p(x) = f_1(x)f_2(x) \ldots f_n(x),$$

avem

$$D(\log\ p(x)) = D(\log f_1(x)) + D(\log f_2(x)) + \ldots + D(\log f_n(x)),$$

şi deci din regula de derivare a funcţiilor compuse

$$\frac{p'(x)}{p(x)} = \frac{f_1'(x)}{f_1(x)} + \frac{f_2'(x)}{f_2(x)} + \ldots + \frac{f_n'(x)}{f_n(x)}.$$

Folosind această formulă, derivaţi funcţiile

a) $x(x + 1)(x + 2) \ldots (x + n)$, b) xe^{-ax^2}.

5. Serii infinite pentru logaritm. Calculul numeric

Pentru calculul numeric al logaritmului nu se foloseşte formula (15). O expresie explicită, dar mai folositoare, cu totul diferită şi de mare importanţă teoretică este mult mai potrivită acestui scop. Vom obţine această expresie cu ajutorul metodei folosite la p. 460 pentru găsirea lui π, folosind definiţia logaritmului prin formula (1). Mai este necesar un mic pas pregătitor; în loc să ne concentrăm atenţia asupra lui $\log x$, vom încerca să exprimăm funcţia $y = \log (1 + x)$, compusă din funcţiile $y = \log z$ şi $z = 1 + x$. Avem $\frac{dy}{dx} = \frac{dy}{dz} \cdot \frac{dz}{dx} = \frac{1}{z} \cdot 1 = \frac{1}{1 + x}$. Deci $\log (1 + x)$ este o funcţie primitivă a funcţiei $1/(1 + x)$ şi din teorema fundamentală deducem că integrala funcţiei $1/(1 + u)$ de la 0 la x este egală cu $\log (1 + x) - \log 1 = \log (1 + x)$; simbolic

$$\log\ (1 + x) = \int_0^x \frac{1}{1 + u}\, du. \tag{16}$$

(Desigur, această formulă ar fi putut fi obţinută intuitiv tot atît de bine din interpretarea geometrică a logaritmului ca arie, cf. p. 431.)

În formula (16), ca la p. 460, să înlocuim funcția $(1 + u)^{-1}$, prin

$$\frac{1}{1+u} = 1 - u + u^2 - u^3 + \ldots + (-1)^{n-1} u^{n-1} + (-1)^n \frac{u^n}{1+u},$$

unde am avut grijă să folosim o sumă finită și nu o serie infinită, dar în care mai intervine și restul

$$R_n = (-1)^n \frac{u^n}{1+u}.$$

Făcînd această substituție în (16), putem folosi regula că o astfel de sumă (finită) poate fi integrată termen cu termen. Integrarea lui u^s de la 0 la x dă $x^{s+1}/(s+1)$ și obținem astfel imediat

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^{n-1} \frac{x^n}{n} + T_n,$$

unde T_n este dat de

$$T_n = (-1)^n \int_0^x \frac{u^n}{1+u}\, du.$$

Vom arăta acum că T_n tinde spre zero cînd n crește, cu condiția ca x să fie ales mai mare decît -1 și cel mult egal cu $+1$; cu alte cuvinte, pentru

$$1 < x \leqslant 1,$$

(remarcăm că $x = +1$, este o valoare admisă, în timp ce $x = -1$ nu este admisă). Conform acestei ipoteze, în intervalul de integrare u este mai mare decît un număr $-\alpha$, care poate fi apropiat de -1, însă în orice caz este mai mare decît -1, astfel încît $0 < 1 - \alpha < 1 + u$. Deci în intervalul de la 0 la x avem

$$\left| \frac{u^n}{1+u} \right| \leqslant \frac{|u|^n}{1-\alpha}$$

și, prin urmare,

$$|T_n| \leqslant \frac{1}{1-\alpha} \left| \int_0^x u^n du \right|,$$

sau

$$|T_n| \leqslant \frac{1}{1-\alpha} \frac{|x|^{n+1}}{n+1} \leqslant \frac{1}{1-\alpha} \frac{1}{n+1}.$$

Deoarece $1 - \alpha$ este un factor fixat, vedem că atunci cînd n crește, această expresie tinde spre 0, astfel încît din

$$(17)\quad \left|\log(1+x) - \left\{x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + (-1)^n \frac{x^n}{n}\right\}\right| \leqslant \frac{1}{1-\alpha}\,\frac{1}{n+1},$$

obținem seria infinită

$$(18)\quad \log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots,$$

care este valabilă pentru $-1 < x \leqslant 1$.

Dacă alegem, în particular, $x = 1$, obținem rezultatul interesant

$$(19)\quad \log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots .$$

Această formulă are o structură asemănătoare cu aceea a seriei care dă pe $\pi/4$.

Seria (18) nu are o însemnătate practică prea mare pentru găsirea valorilor numerice ale logaritmului, deoarece domeniul de variație al cantității $1 + x$ se limitează la intervalul de la 0 la 2 și deoarece convergența ei este atît de înceată, încît trebuie să luăm mulți termeni pentru a obține un rezultat destul de precis. Prin artificiul următor putem obține o expresie mai convenabilă. Înlocuind pe x cu $-x$ în (18), găsim

$$(20)\quad \log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \ldots$$

Mai departe, scăzînd formula (20) din formula (18) și folosind faptul că $\log a - \log b = \log a + \log(1/b) = \log(a/b)$, obținem

$$(21)\quad \log\frac{1+x}{1-x} = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \ldots\right).$$

Nu numai că această serie converge mult mai repede, dar membrul stîng al formulei poate exprima acum logaritmul oricărui număr pozitiv z, deoarece $\frac{1+x}{1-x} = z$ are întotdeauna o soluție x cuprinsă între -1 și $+1$. Astfel, dacă vrem să calculăm $\log 3$, punem $x = 1/2$ și obținem

$$\log 3 = \log\frac{1+\frac{1}{2}}{1-\frac{1}{2}} = 2\left(\frac{1}{1\cdot 2} + \frac{1}{3\cdot 2^3} + \frac{1}{5\cdot 2^5} + \ldots\right).$$

Luînd doar şase termeni, pînă la $\frac{2}{11 \cdot 2^{11}} = \frac{1}{11264}$, găsim valoarea cu cinci cifre exacte: $\log 3 = 1{,}0986$.

§ 7. ECUAŢII DIFERENŢIALE

1. Definiţie

Rolul dominant al funcţiilor exponenţiale şi trigonometrice din analiza matematică şi aplicaţiile ei la problemele fizice se explică prin faptul că aceste funcţii rezolvă cele mai simple „ecuaţii diferenţiale".

O ecuaţie diferenţială cu o funcţie necunoscută $u = f(x)$ cu derivata $u' = f'(x)$ — notaţia u' este o prescurtare foarte folositoare pentru $f'(x)$, atîta timp cît cantitatea u şi dependenţa ei de x ca funcţie $f(x)$ nu trebuie să fie scoasă în evidenţă — este o ecuaţie în care intervin u, u' şi eventual variabila independentă x ca, de pildă,

$$u' = u + \sin(xu)$$

sau

$$u' + 3u = x^2.$$

Mai general, o ecuaţie diferenţială poate conţine derivata de ordinul doi $u'' = f''(x)$ sau derivate de ordin superior, ca în exemplul

$$u'' + 2u' - 3u = 0.$$

În toate cazurile, problema constă în găsirea unei funcţii $u = f(x)$, care satisface ecuaţia dată. Rezolvarea unei ecuaţii diferenţiale este o generalizare a problemei integrării, în sensul găsirii funcţiei primitive a unei funcţii date $g(x)$, care se reduce la rezolvarea ecuaţiei diferenţiale simple

$$u' = g(x).$$

De exemplu, soluţiile ecuaţiei diferenţiale

$$u' = x^2$$

sînt funcţiile $u = x^3/3 + c$, unde c este o constantă arbitrară.

2. Ecuaţia diferenţială a funcţiei exponenţiale. Dezintegrarea radioactivă. Legea creşterii. Dobînda compusă

Ecuaţia diferenţială

(1) $$u' = u$$

are ca soluţie funcţia exponenţială $u = e^x$, deoarece funcţia exponenţială este egală cu derivata ei. Mai general, funcţia $u = ce^x$, unde c este o constantă arbitrară, este o soluţie a ecuaţiei (1). În mod analog, funcţia

(2) $$u = ce^{kx}.$$

unde c şi k sînt două constante arbitrare, este o soluţie a ecuaţiei diferenţiale

$$u' = ku. \tag{3}$$

Reciproc, orice funcţie $u = f(x)$ care satisface ecuaţia (3) trebuie să fie de forma ce^{kx}, pentru că dacă $x = h(u)$ este funcţia inversă a funcţiei $u = f(x)$, atunci conform cu regula de calcul a derivatei unei funcţii inverse, avem:

$$h' = \frac{1}{u'} = \frac{1}{ku}.$$

Însă $\frac{\log u}{k}$ este o primitivă a funcţiei $\frac{1}{ku}$, astfel încît $x = h(u) = \frac{1}{k} \log u + b$, unde b este o constantă. Deci

$$\log u = kx - bk,$$

şi

$$u = e^{kx} \cdot e^{-bk}.$$

Egalînd pe e^{-bk} (care este o constantă) cu c, avem

$$u = ce^{kx},$$

ceea ce trebuia demonstrat.

Importanţa ecuaţiei diferenţiale (3) constă în faptul că ea guvernează procese fizice, în care o cantitate u de substanţă depinde de timpul t,

$$u = f(t),$$

şi în care cantitatea u variază în fiecare moment, cu o viteză proporţională cu cantitatea u existentă în acel moment. În acest caz, *viteza de variaţie* la momentul t,

$$u' = f'(t) = \lim \frac{f(t_1) - f(t)}{t_1 - t},$$

este egală cu ku, unde k este o constantă pozitivă, dacă u creşte, şi negativă, dacă u descreşte. În orice caz, u satisface ecuaţia diferenţială (3); deci

$$u = ce^{kt}.$$

Constanta c este determinată dacă cunoaştem cantitatea u_0 existentă la momentul $t = 0$. Trebuie să obţinem această cantitate dacă punem $t = 0$ în ecuaţia (2):

$$u_0 = ce^0 = c,$$

astfel încît

(4) $$u = u_0 e^{kt}.$$

Remarcăm că pornim de la ipoteza cunoaşterii *vitezei de variaţie* a lui u şi deducem legea (4), care dă *cantitatea* reală u la orice moment t. Aceasta este chiar inversa problemei găsirii derivatei unei funcţii.

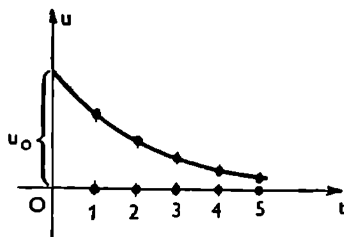


Fig. 279. Descreşterea exponenţială; $u = u_0 e^{kt}$, $k < 0$

Un exemplu tipic este acela al dezintegrării radioactive. Fie $u = f(t)$ cantitatea de substanţă radioactivă la momentul t; atunci, în ipoteza că fiecare particulă individuală de substanţă are o anumită probabilitate de a se dezintegra într-un anumit interval de timp şi că probabilitatea nu este modificată prin prezenţa altor particule de acelaşi fel, viteza de dezintegrare a lui u la un anumit moment t va fi proporţională cu u, adică cu cantitatea totală existentă în acel moment. Deci u va satisface ecuaţia (3) cu o constantă negativă k, care măsoară viteza procesului de dezintegrare şi, de aceea,

$$u = u_0 e^{kt}.$$

Rezultă că fracţiunea din u care se dezintegrează în două intervale egale de timp este aceeaşi, pentru că dacă u_1 este cantitatea existentă la momentul t_1 iar u_2 este cantitatea existentă într-un alt moment t_2, ulterior lui t_1, atunci:

$$\frac{u_2}{u_1} = \frac{u_0 e^{kt_2}}{u_0 e^{kt_1}} = e^{k(t_2 - t_1)},$$

care depinde numai de $t_2 - t_1$. Pentru a găsi cît timp este necesar pentru ca o anumită cantitate de substanţă să se dezintegreze pînă ce rămîne o jumătate din cantitatea iniţială, trebuie să-l determinăm pe $s = t_2 - t_1$, astfel încît

$$\frac{u_2}{u_1} = \frac{1}{2} = e^{ks},$$

de unde găsim

$$(5) \qquad ks = \log \frac{1}{2}, \ s = (-\log 2)/k \text{ sau } k = (-\log 2)/s.$$

Pentru orice substanță radioactivă, valoarea lui s se numește timpul de înjumătățire, iar s, sau o valoare asemănătoare (ca de pildă valoarea r, pentru care $u_2/u_1 = 999/1000$) poate fi găsită pe cale experimentală. Pentru radiu, timpul de înjumătățire este de aproximativ 1 550 de ani și

$$k = \frac{\log \frac{1}{2}}{1\,550} = -0{,}0000447.$$

Rezultă că

$$u = u_0 \cdot e^{-0{,}0000447t}.$$

Un exemplu de lege de creștere care este aproximativ exponențială este dat de fenomenul dobînzii compuse. O anumită sumă de bani, de exemplu u_0 dolari, este depusă cu o dobîndă compusă de 3%, care trebuie adăugată în fiecare an. După un an, suma de bani, va fi de

$$u_1 = u_0(1 + 0{,}03),$$

după doi ani va fi de

$$u_2 = u_1(1 + 0{,}03) = u_0(1 + 0{,}03)^2,$$

iar după t ani va fi de

$$u_t = u_0(1 + 0{,}03)^t.$$

Acum, dacă în loc de a adăuga dobînda la sfîrșitul fiecărui an, o adăugăm la sfîrșitul fiecărei luni, sau după fiecare a n-a parte dintr-un an, atunci după t ani suma acumulată va fi:

$$u_0\left(1 + \frac{0{,}03}{n}\right)^{nt} = u_0\left[\left(1 + \frac{0{,}03}{n}\right)^n\right]^t.$$

Dacă n este foarte mare, astfel încît dobînda să fie adăugată zilnic sau din oră în oră, atunci cînd n tinde spre infinit, cantitatea din paranteze le drepte conform cu § 6, tinde spre $e^{0{,}03}$. Deci la limită, după t ani, suma totală va fi de

$$(7) \qquad u_0 \cdot e^{0{,}03t},$$

care corespunde unui proces continuu de dobîndă compusă. Putem calcula, de asemenea, timpul s, necesar pentru dublarea capitalului inițial, cu dobînda

compusă de 3%. Avem $\frac{u_0 \cdot e^{0,03s}}{u_0} = 2$, astfel încît $s = \frac{100}{3} \log 2 = 23,10$.
Astfel suma de bani se dublează după aproximativ 23 de ani.

În loc de a urma acest procedeu treptat și de a trece apoi la limită, am fi putut obține formula (7), spunînd pur și simplu că viteza de creștere u' a capitalului este proporțională cu u, cu factorul $k = 0,03$, astfel încît

$$u' = ku \quad \text{unde} \quad k = 0,03.$$

Formula (7) se obține atunci din rezultatul general (4).

3. Alte exemple. Cele mai simple vibrații

Funcția exponențială apare adesea în combinații mai complicate. De exemplu, funcția

$$u = e^{-kx^2}, \tag{8}$$

unde k este o constantă pozitivă, este o soluție a ecuației diferențiale

$$u' = -2kxu.$$

Funcția (8) are o importanță fundamentală în teoria probabilităților și statistică, deoarece ea definește distribuțiile de frecvență „normală".

Funcțiile trigonometrice $u = \cos t$, $v = \sin t$ satisfac, de asemenea, o ecuație diferențială foarte simplă. Avem mai întîi

$$u' = -\sin t = -v.$$

$$v' = \cos t = u,$$

care este un „sistem de două ecuații diferențiale cu două funcții necunoscute". Derivînd din nou, găsim

$$u'' = -v' = -u,$$

$$v'' = u' = -v,$$

astfel încît cele două funcții u și v de variabila temporală t pot fi considerate drept soluții ale aceleiași ecuații diferențiale

$$z'' + z = 0, \tag{9}$$

care este o ecuație diferențială foarte simplă de „ordinul doi", adică în ea intervine derivata de ordinul doi a lui z. Această ecuație, sau mai bine spus generalizarea ei, care conține o constantă pozitivă k^2

$$z'' + k^2 z = 0, \tag{10}$$

ale cărei soluții sînt $z = \cos kt$ și $z = \sin kt$, intervine adesea în studiul teoriei vibrațiilor. Acesta este motivul pentru care curbele oscilante $u = \sin kt$ și $u = \cos kt$ (fig. 280) formează scheletul teoriei vibrațiilor. Trebuie să observăm că ecuația diferențială (10) reprezintă cazul ideal, cînd se presupune că lipsește frecarea, sau rezistența. Rezistența apare în ecuația diferențială a vibrațiilor printr-un alt termen, rz'

11) $$z'' + rz' + k^2z = 0,$$

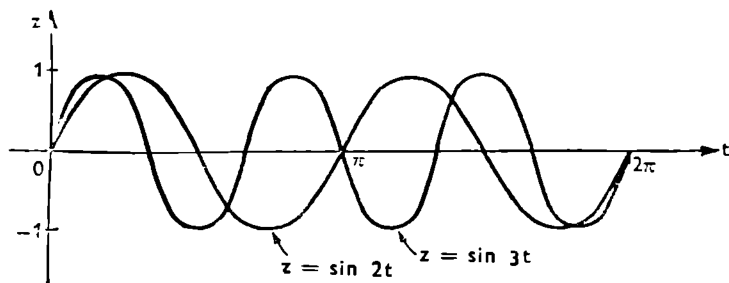


Fig. 280.

iar soluțiile sînt acum vibrații „amortizate", exprimate matematic prin formula

$$e^{-rt/2} \cos \omega t, \quad e^{-rt/2} \sin \omega t; \; \omega = \sqrt{k^2 - \left(\frac{r}{2}\right)^2},$$

și reprezentate grafic în fig. 281. (Ca exercițiu, cititorul poate verifica faptul că aceste funcții sînt soluții, efectuînd derivările.) Oscilațiile de aici sînt de același tip ca și acelea ale sinusului sau cosinusului obișnuit, dar cu timpul ele sînt amortizate în intensitate printr-un factor exponențial, care descrește mai repede sau mai încet, în funcție de mărimea coeficientului de frecare r.

4. Legea lui Newton a dinamicii

Cu toate că o analiză mai detaliată a acestor fapte depășește scopul nostru, vrem să le dăm o formă generală în cadrul conceptelor fundamentale prin care Newton a revoluționat mecanica și fizica. El a considerat mișcarea unei particule de masă m și de coordonate spațiale $x(t)$, $y(t)$, $z(t)$, care sînt funcții de timpul t, astfel încît componentele accelerației sînt derivatele de ordinul doi $x''(t)$, $y''(t)$, $z''(t)$. Pasul cel mai important făcut de Newton a fost recunoașterea faptului că cantitățile mx'', my'', mz'' pot fi considerate drept componente ale forței care acționează asupra particulei. La prima vedere,

acest lucru ar putea să pară a fi doar o definiţie formală a cuvîntului ,,forţă" din fizică. Dar marele merit al lui Newton a fost acela că el a formulat această definiţie în concordanţă cu fenomenele naturale reale, deoarece natura creează foarte des cîmpuri de astfel de forţe, care ne sînt cunoscute, fără a cunoaşte mişcarea particulară pe care vrem s-o studiem.

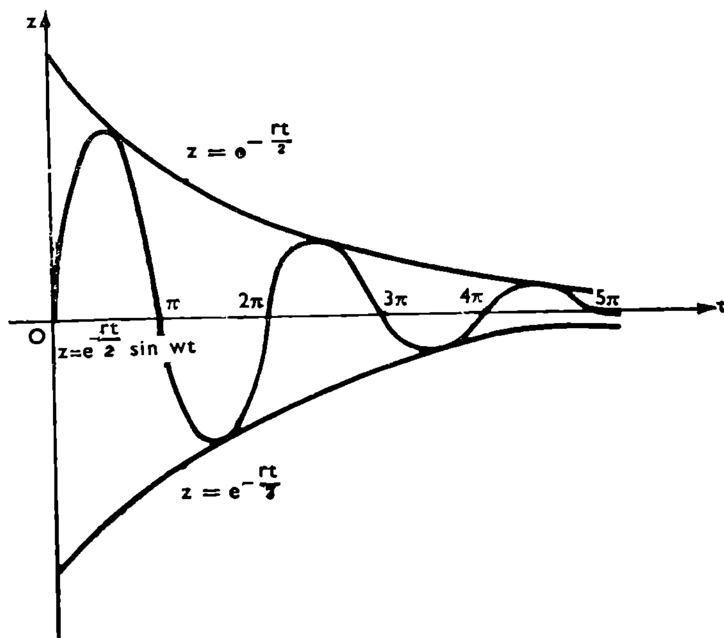


Fig. 281. Vibraţii amortizate

Marele triumf al lui Newton în dinamică, justificarea legii lui Kepler a mişcării planetelor, demonstrează clar concordanţa care există între concepţia sa matematică şi fenomenele naturii. Newton a presupus la început că atracţia gravitaţiei este invers proporţională cu pătratul distanţei. Dacă aşezăm Soarele în originea sistemului de coordonate şi dacă o planetă dată are coordonatele x, y, z, atunci rezultă că componentele forţei în direcţiile Ox, Oy, Oz sînt respectiv egale cu

$$-k \cdot \frac{x}{r^3}; \qquad -k \cdot \frac{y}{r^3}, \qquad -k \cdot \frac{z}{r^3},$$

unde k este o constantă gravitaţională care nu depinde de timp, iar $r = \sqrt{x^2 + y^2 + z^2}$ este distanţa de la Soare la planetă. Aceste expresii deter-

mină cîmpul local de forțe, independent de mișcarea particulei în cîmp. Acum, această cunoaștere a cîmpului de forțe este combinată cu legea generală a dinamicii a lui Newton (adică, legătura dintre forță și accelerație); din egalarea celor două expresii rezultă ecuațiile

$$mx'' = \frac{-kx}{(x^2+y^2+z^2)^{3/2}}, \quad my'' = \frac{-ky}{(x^2+y^2+z^2)^{3/2}},$$

$$mz'' = \frac{-kz}{(x^2+y^2+z^2)^{3/2}},$$

un sistem format din trei ecuații diferențiale cu trei funcții necunoscute $x(t)$, $y(t)$, $z(t)$. Acest sistem poate fi rezolvat și rezultă că, în concordanță cu observațiile empirice ale lui Kepler, orbita planetei este o secțiune conică care are Soarele într-un focar, ariile descrise de un segment care unește Soarele cu planeta sînt egale în intervale de timp egale și pătratele perioadelor de revoluție a două planete sînt proporționale cu cuburile distanțelor lor pînă la Soare. Vom omite demonstrația acestui fapt.

Problema vibrațiilor dă o ilustrare mai elementară a metodei lui Newton. Să presupunem că avem o particulă care se mișcă de-a lungul unei drepte, axa Ox, și este legată de origine printr-o forță elastică, ca de pildă un resort sau un fir de cauciuc. Dacă particula este scoasă din poziția ei de echilibru (din origine) și se plasează într-o poziție de coordonată x, atunci forța o va atrage înapoi cu o intensitate pe care o presupunem proporțională cu depărtarea x; deoarece forța este îndreptată spre origine, ea va fi reprezentată de $-k^2x$, unde $-k^2$ este un factor negativ de proporționalitate, care exprimă rezistența resortului elastic sau a firului de cauciuc. Mai mult, să presupunem că există și o frecare care întîrzie mișcarea și că această frecare este proporțională cu viteza x' a particulei, cu un factor de proporționalitate $-r$. Atunci, forța totală în orice moment va fi dată de $-k^2x - rx'$ și, conform principiului general al lui Newton, găsim $mx'' = -k^2x - rx'$ sau

$$mx'' + rx' + k^2x = 0.$$

Aceasta este chiar ecuația diferențială (11) a vibrațiilor amortizate, menționată mai sus.

Acest exemplu este de mare importanță, deoarece numeroase tipuri de sisteme mecanice și electrice oscilante pot fi descrise matematic chiar prin această ecuație diferențială. Aici avem un exemplu tipic în care o formulare matematică abstractă dezvăluie structura internă a numeroase fenomene în aparență total diferite și fără vre-o legătură între ele. Această abstracție de la natura particulară a unui fenomen dat la formularea legii generale care guvernează o întreagă clasă de fenomene este una dintre trăsăturile caracteristice ale tratării matematice a problemelor fizicii.

§ 1. PROBLEME DE PRINCIPIU

1. Derivabilitatea

Am legat noțiunea de derivată a unei funcții $y = f(x)$ de ideea intuitivă a tangentei la graficul unei funcții. Deoarece noțiunea generală de funcție este atît de largă, pentru o expunere logică completă trebuie să renunțăm la această dependență de intuiția geometrică. Într-adevăr, nu avem nici o garanție că faptele intuitive cunoscute din considerarea unor curbe simple, ca de pildă cercuri și elipse, vor rămîne în vigoare pentru graficele unor funcții mai complicate. Să considerăm, de exemplu, funcția din fig. 282, al cărei grafic are un vîrf. Această funcție este definită de ecuația $y = x + |x|$, unde $|x|$ este valoarea absolută a lui x, adică

$$y = x + x = 2x \qquad \text{pentru } x \geqslant 0,$$

$$y = x - x = 0 \qquad \text{pentru } x < 0.$$

Un alt exemplu de același fel este funcția $y = |x|$; un altul este funcția $y = x + |x| + (x - 1) + |x - 1|$. Graficele acestor funcții nu au tangentă definită în anumite puncte; aceasta înseamnă că funcțiile nu au derivate pentru valorile corespunzătoare ale lui x.

Exerciții: 1) Construiți funcția $f(x)$ al cărei grafic este o jumătate de hexagon regulat.

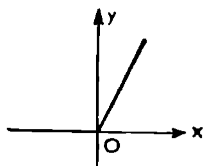


Fig. 282. $y = x + |x|$

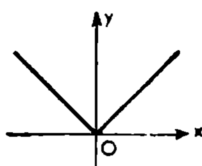


Fig. 283. $y = |x|$

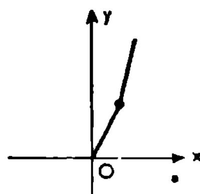


Fig. 284. $y = x + |x| + (x - 1) + |x - 1|$

2) Unde sînt vîrfurile graficului funcției

$$f(x) = (x + |x|) + \frac{1}{2}\left\{\left(x - \frac{1}{2}\right) + \left|x - \frac{1}{2}\right|\right\} + \frac{1}{4}\left\{\left(x - \frac{1}{4}\right) + \left|x - \frac{1}{4}\right|\right\} ?$$

Care sînt discontinuităţile lui $f'(x)$?

Ca alt exemplu de simplu nederivabilitate să considerăm funcţia

$$y = f(x) = x \sin \frac{1}{x},$$

care se obţine din funcţia $\sin(1/x)$ (cf. p. 300), înmulţind-o cu factorul x; definim pe $f(x)$ ca fiind nulă pentru $x = 0$. Această funcţie, al cărei grafic pentru valorile pozitive ale lui x este arătat în fig. 285, este continuă peste tot.

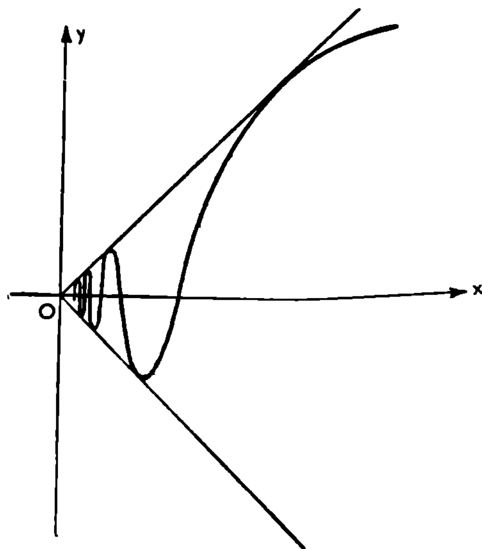


Fig. 285. y = x *sin* 1/x.

Graficul oscilează de o infinitate de ori în vecinătatea lui $x = 0$, ,,undele'' devenind foarte mici pe măsură ce ne apropiem de $x = 0$. Panta acestor unde este dată de

$$f'(x) = \sin \frac{1}{x} - \frac{1}{x} \cos \frac{1}{x}$$

(cititorul poate verifica acest lucru ca exerciţiu); cînd x tinde spre 0, această pantă oscilează între margini pozitive şi negative, care cresc mereu. Pentru

$x = 0$ putem încerca să găsim derivata, ca limită pentru $h \to 0$ a cîtului creşterilor

$$\frac{f(0+h)-f(0)}{h} = \frac{h \sin \frac{1}{h}}{h} = \sin \frac{1}{h},$$

dar cînd $h \to 0$, acest cît al creşterilor oscilează între -1 şi $+1$ şi nu tinde spre o limită; deci funcţia nu poate fi derivată în $x = 0$.

Aceste exemple indică o dificultate inerentă acestui subiect. Weierstrass a ilustrat situaţia în modul cel mai frapant, construind o funcţie continuă, al cărei grafic nu are tangentă în nici un punct. În timp ce derivabilitatea implică continuitatea, acest fapt arată că continuitatea nu implică derivabilitatea, deoarece funcţia lui Weierstrass este continuă şi nederivabilă în nici un punct. În practică nu apar astfel de dificultăţi. Cu excepţia eventual a unor puncte izolate, curbele vor fi netede iar derivarea nu va fi numai posibilă, dar va da şi o derivată continuă. De ce nu cerem atunci ca fenomenele ,,patologice'' să lipsească din problemele considerate? Tocmai acest lucru facem în calculul diferenţial, cînd considerăm numai funcţii derivabile. În cap. VIII am efectuat derivarea unei clase largi de funcţii, demonstrînd în acest mod derivabilitatea lor.

Deoarece derivabilitatea unei funcţii nu este un fapt logic de la sine înţeles, ea trebuie să fie presupusă sau demonstrată. Noţiunea de tangentă sau de direcţie a unei curbe, care iniţial a fost baza noţiunii de derivată, se deduce atunci din definiţia pur analitică a derivatei; dacă funcţia $y = f(x)$ are o derivată, adică dacă cîtul creşterilor $\frac{f(x+h)-f(x)}{h}$ are o limită unică $f'(x)$, cînd h tinde spre 0 din ambele părţi, atunci se spune că curba corespunzătoare are o tangentă cu panta $f'(x)$. În acest mod, atitudinea naivă a lui Fermat, Leibniz şi Newton este inversată în interesul corectitudinii logice.

Exerciţii: 1) Arătaţi că funcţia continuă definită de $x^2 \sin(1/x)$ are derivată în $x = 0$.

2) Arătaţi că funcţia arc tg$(1/x)$ este discontinuă pentru $x = 0$, că x arc tg$(1/x)$ este continuă, dar nu are derivată, şi că x^2 arc tg$(1/x)$ are derivată în $x = 0$.

2. Integrala

Situaţia este asemănătoare în privinţa integralei unei funcţii continue $f(x)$. În loc de a considera ,,aria mărginită de curba'' $y = f(x)$ ca o cantitate a cărei existenţă este evidentă şi care poate fi exprimată *a posteriori*, ca limită a unei sume, *definim* integrala prin această limită şi considerăm integrala ca noţiune de bază din care deducem apoi noţiunea generală de arie. Această

atitudine ne este impusă atuuci cînd recunoaştem imprecizia intuiției geometrice aplicată unor noțiuni analitice atît de generale, ca de pildă aceea de funcție continuă. Începem cu formarea unei sume

$$(1) \qquad S_n = \sum_{j=1}^{n} f(v_j)(x_j - x_{j-1}) = \sum_{j=1}^{n} f(v_j)\Delta x_j,$$

unde $x_0 = a, x_1, \ldots, x_n = b$ este o diviziune a intervalului de integrare, $\Delta x_j = x_j - x_{j-1}$ este lungimea celui de al j-lea interval parțial, iar v_j este o valoare arbitrară a lui x din acest interval parțial, adică $x_{j-1} \leqslant v_j \leqslant x_j$. (Putem lua, de exemplu, $v_j = x_j$ sau $v_j = x_{j-1}$.) Mai departe, formăm un şir de astfel de sume, în care numărul n al subintervalelor creşte şi, în acelaşi timp, lungimea maximă a intervalelor parțiale descreşte spre zero. Atunci rezultatul principal este următorul: pentru o funcție continuă dată $f(x)$, suma S_n tinde spre o anumită limită A, independentă de modul particular în care sînt alese intervalele parțiale şi puuctele v_j. Prin definiție, această limită este integrala $A = \int_a^b f(x)dx$. Desigur, existența acestei limite are nevoie de o demonstrație analitică, dacă nu vrem să ne bazăm pe noțiunea geometrică intuitivă de arie. Această demonstrație este dată în orice manual riguros de analiză.

Comparînd derivarea şi integrarea, sîntem puşi în fața următoarei situații antitetice. Derivabilitatea este o condiție restrictivă impusă unei funcții continue, însă efectuarea derivării, adică algoritmul calculului diferențial, este în practică un procedeu direct, bazat pe cîteva reguli simple. Pe de altă parte, orice funcție continuă, fără excepție, are o integrală între orice două limite date. Însă calculul explicit al acestor integrale, chiar pentru funcțiile foarte simple, este în general o problemă foarte dificilă. În acest moment, teorema fundamentală a analizei devine în multe cazuri instrumentul decisiv de efectuare a integrării. Şi cu toate acestea, pentru cele mai multe funcții, chiar pentru unele foarte elementare, integrarea nu dă expresii explicite simple, iar calculul numeric al integralelor necesită metode avansate.

3. Alte aplicații ale noțiunii de integrală. Lucru mecanic. Lungime

Separînd noțiunea analitică de integrală de interpretarea ei geometrică inițială, găsim alte interpretări şi aplicații tot atît de importante. De exemplu, integrala poate fi utilizată în mecanică pentru exprimarea noțiunii de lucru mecanic. Următorul caz foarte simplu va fi suficient pentru explicația noastră. Să presupunem că un punct material se mişcă de-a lungul axei Ox, sub influența

unei forțe dirijate de-a lungul axei. Această masă este gîndită ca fiind concentrată în punctul de coordonată x, iar forța este dată ca funcție $f(x)$ de poziția punctului, semnul lui $f(x)$ indicînd dacă ea este îndreptată în sensul pozitiv sau în cel negativ al axei Ox. Dacă forța este constantă și deplasează punctul material de la a la b, atunci lucrul mecanic efectuat este dat de produsul $(b - a)f$ dintre intensitatea f a forței și distanța parcursă de punct. Dacă însă intensitatea variază o dată cu x, va trebui să definim cantitatea de lucru mecanic efectuat, printr-un proces de trecere la limită (tot așa cum am definit și viteza). În acest scop, împărțim intervalul de la a la b, ca și mai înainte în intervale parțiale mici, prin puncele $x_0 = a, x_1, \ldots, x_n = b$; apoi „ne închipuim că, în fiecare interval parțial, forța este constantă și egală, de pildă cu $f(x_\nu)$, valoarea reală la o extremitate a intervalului. Lucrul mecanic corespunzător acestei forțe, care variază în salturi, este dat de expresia

$$S_n = \sum_{\nu=1}^{n} f(x_\nu)\Delta x_\nu .$$

Dacă rafinăm acum diviziunea ca mai înainte, și facem pe n să crească, vedem că suma tinde spre integrala

$$\int_a^b f(x)\,\mathrm{d}x.$$

Astfel, lucrul mecanic efectuat de o forță care variază continuu este definit printr-o integrală.

Ca exemplu, să considerăm o masă m, legată de originea $x = 0$, printr-un resort elastic. Conform celor spuse la p. 480, forța $f(x)$ va fi proporțională cu x:

$$f(x) = -k^2x,$$

unde k^2 este o constantă pozitivă. Atunci lucrul mecanic efectuat de această forță, dacă punctul material se deplasează din origine în poziția $x = b$, va fi

$$\int_0^b -k^2x\mathrm{d}x = -k^2\frac{b^2}{2},$$

și lucrul mecanic pe care trebuie să-l efectuăm pentru a învinge această forță, dacă vrem să întindem resortul pînă în poziția $x = b$, este egal cu $+k^2\frac{b^2}{2}$.

O altă aplicație a noțiunii generale de integrală se referă la noțiunea de lungime a unui arc de curbă. Să presupunem că porțiunea de curbă considerată este reprezentată printr-o funcție $y = f(x)$, a cărei derivată $f'(x) = \mathrm{d}y/\mathrm{d}x$

este şi ea continuă. Pentru a defini lungimea, procedăm ca şi cum ar trebui să măsurăm curba în scopuri practice, cu ajutorul unei rigle. Înscriem în arcul AB un polinom cu n laturi mici, măsurăm lungimea totală L_n a acestui poligon şi considerăm lungimea L_n ca fiind o aproximaţie; făcîndu-l pe n să crească, astfel încît lungimea maximă a laturilor poligonului să descrească spre zero, definim numărul

$$L = \lim L_n$$

ca fiind lungimea arcului AB. (În cap. VI lungimea unui cerc a fost obţinută în acest mod, ca limită a perimetrelor n-goanelor regulate înscrise.) Se poate arăta că pentru curbe destul de netede, limita există şi este independentă de modul particular în care se alege şirul de poligoane înscrise. Curbele pentru care se întîmplă acest lucru se numesc *rectificabile*. Orice curbă ,,rezonabilă" care apare în teorie sau aplicaţii va fi rectificabilă şi nu vom insista cu studiul cazurilor patologice. Va fi suficient să arătăm că arcul AB, pentru o funcţie $y = f(x)$, cu derivata $f'(x)$ continuă, are o lungime L în acest sens şi că L poate fi exprimată printr-o integrală.

În acest scop să notăm abscisele lui A şi B, respectiv cu a şi b şi să împărţim intervalul cuprins între a şi b, ca mai înainte, prin punctele $x_0 = a$, $x_1, \ldots, x_j, \ldots x_n = b$, cu diferenţele $\Delta x_j = x_j - \Delta_{j-1}$ şi să considerăm poligonul cu vîrfurile de coordonate x_j, $y_j = f(x_j)$, aflate deasupra acestor puncte de diviziune. O latură a poligonului va avea lungimea

$$\sqrt{(x_j - x_{j-1})^2 + (y_j - y_{j-1})^2} = \sqrt{(\Delta x_j)^2 + (\Delta y_j)^2} = \Delta x_j \sqrt{1 + \left(\frac{\Delta y_j}{\Delta x_j}\right)^2}.$$

Deci pentru lungimea totală a poligonului avem formula

$$L_n = \sum_{j=1}^{n} \sqrt{1 + \left(\frac{\Delta y_j}{\Delta x_j}\right)^2}\, \Delta x_j.$$

Dacă acum n tinde spre infinit, cîturile creşterilor $\frac{\Delta y_j}{\Delta x_j}$ vor tinde spre derivata $\frac{dy}{dx} = f'(x)$ şi pentru lungimea L obţinem o expresie sub forma unei integrale

$$L = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx.$$

Fără a intra mai mult în amănuntele acestei discuţii teoretice, mai facem două observaţii suplimentare. În primul rînd, dacă punctul B este considerat

variabil pe curbă cu coordonata x, atunci $L = L(x)$ este o funcție de x și din teorema fundamentală avem

$$L'(x) = \frac{dL}{dx} = \sqrt{1 + [f'(x)]^2},$$

o formulă folosită în mod frecvent. În al doilea rînd, în timp ce formula (2) dă soluția „generală" a problemei, în cazurile particulare ea dă rareori o expresie explicită a lungimii arcului. În acest scop, trebuie să substituim

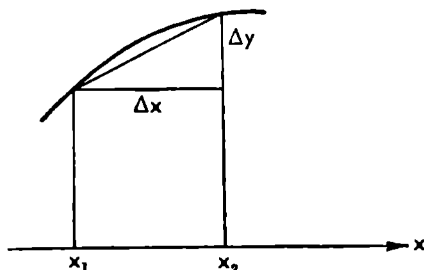


Fig. 286. Lungimea unui arc

funcția particulară $f(x)$ sau mai degrabă $f'(x)$, în formula (2), și apoi trebuie să efectuăm integrarea expresiei obținute. Aici apare, în general, o dificultate de neînlăturat, dacă ne restrîngem la domeniul funcțiilor elementare, considerate în această carte. Vom menționa cîteva cazuri, în care integrarea este posibilă. Funcția

$$y = f(x) = \sqrt{1 - x^2}$$

reprezintă cercul unitate; avem $f'(x) = \frac{dy}{dx} = -\frac{x}{\sqrt{1-x^2}}$, de unde $\sqrt{1 + f'(x)^2} = \frac{1}{\sqrt{1-x^2}}$; prin urmare, lungimea arcului de cerc este dată de integrala

$$\int_a^b \frac{dx}{\sqrt{1-x^2}} = \operatorname{arc\,sin} b - \operatorname{arc\,sin} a.$$

Pentru parabola $y = x^2$ avem $f'(x) = 2x$ și lungimea arcului cuprins între $x = 0$ și $x = b$ este egală cu

$$\int_a^b \sqrt{1 + 4x^2}\, dx.$$

Pentru curba $y = \log \sin x$ avem $f'(x) = \operatorname{ctg} x$, iar lungimea arcului se exprimă prin integrala

$$\int_a^b \sqrt{1 + \operatorname{ctg}^2 x}\, dx.$$

Ne vom mulţumi doar să scriem aceste expresii integrale. Ele ar putea fi calculate, cunoscînd puţin mai multă tehnică decît dispunem pînă în prezent, dar nu vom merge mai departe în această direcţie.

§ 2. ORDINE DE MĂRIME

1. Funcţia exponenţială şi puterile lui x

În matematică întîlnim adesea şiruri a_n, care tind spre infinit. Deseori trebuie să comparăm un astfel de şir cu un altul b_n, care tinde şi el spre infinit, eventual „mai repede" decît a_n. Pentru a preciza această noţiune, vom spune că b_n tinde spre infinit mai repede decît a_n sau are un *ordin de mărime superior* aceluia al lui a_n, dacă cîtul a_n/b_n (în care atît numărătorul, cît şi numitorul tind spre infinit) tinde spre zero, cînd n creşte. Astfel, şirul $b_n = n^2$ tinde spre infinit mai repede decît şirul $a_n = n$, iar acesta, la rîndul lui, mai repede decît $c_n = \sqrt{n}$, pentru că

$$\frac{a_n}{b_n} = \frac{n}{n^2} = \frac{1}{n} \to 0, \quad \frac{c_n}{a_n} = \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}} \to 0.$$

Este limpede că n^s tinde spre infinit mai repede decît n^r, ori de cîte ori $s > r > 0$, deoarece $n^r/n^s = 1/n^{s-r} \to 0$.

Dacă cîtul a_n/b_n tinde spre o constantă finită c, diferită de zero, spunem că cele două şiruri a_n şi b_n tind spre infinit cu aceeaşi viteză sau au *acelaşi ordin de mărime*. Astfel, de exemplu, $a_n = n^2$ şi $b_n = 2n^2 + n$ au acelaşi ordin de mărime, deoarece

$$\frac{a_n}{b_n} = \frac{n^2}{2n^2 + n} = \frac{1}{2 + \frac{1}{n}} \to \frac{1}{2}.$$

Ne-am putea închipui că, folosind puterile lui n ca diviziunile unei rigle, am putea măsura diferitele viteze de tindere spre infinit pentru orice şir a_n care tinde spre infinit. Pentru a face acest lucru, ar trebui să găsim o putere convenabilă n^s, cu acelaşi ordin de mărime ca şi a_n, adică astfel încît a_n/n^s să tindă spre o constantă finită, diferită de zero. Este un fapt remarcabil că acest lucru nu este întotdeauna posibil, deoarece *funcţia exponenţială*

a^n, *cu* $a > 1$ (*de exemplu* e^n) ***tinde spre infinit mai repede decît orice putere*** n^s, ***oricît de mare l-am alege pe*** s, ***în timp ce*** $\log n$ ***tinde spre infinit mai încet decît orice putere*** n^s, ***oricît de mic ar fi exponentul pozitiv*** s. Cu alte cuvinte, avem relațiile

$$\frac{n^s}{a^n} \to 0 \tag{1}$$

și

$$\frac{\log n}{n^s} \to 0, \tag{2}$$

cînd $n \to \infty$. Nu este necesar ca exponentul s să fie un întreg, ci poate fi orice număr pozitiv fixat.

Pentru a demonstra relația (1), să facem mai întîi o simplificare, extrăgînd rădăcina de ordinul s din cît; dacă rădăcina tinde spre zero, atunci și cîtul inițial are aceeași proprietate. Deci trebuie să demonstrăm doar că

$$\frac{n}{a^{n/s}} \to 0,$$

cînd n crește. Fie $b = a^{1/s}$; deoarece am presupus că a este mai mare decît 1, atunci și b și $\sqrt{b} = b^{1/2}$ vor fi mai mari decît 1. Putem scrie

$$b^{\frac{1}{2}} = 1 + q,$$

unde q este un număr pozitiv. Dacă folosim inegalitatea (6) de la p. 32, avem:

$$b^{n/2} = (1 + q)^n \geqslant 1 + nq > nq,$$

astfel încît

$$a^{n/s} = b^n > n^2q^2$$

și, prin urmare,

$$\frac{n}{a^{n/s}} < \frac{n}{n^2q^2} = \frac{1}{nq^2}.$$

Deoarece ultima cantitate tinde spre zero cînd $n \to \infty$, demonstrația este completă.

De fapt, relația

$$\frac{x^s}{a^x} \to 0 \tag{3}$$

este valabilă cînd x tinde spre infinit în *orice mod*, parcurgînd un şir x_1, $x_2, \ldots$ care nu coincide în mod necesar cu şirul 1, 2, 3, ... al întregilor pozitivi. Într-adevăr, dacă $n - 1 \leqslant x \leqslant n$, atunci:

$$\frac{x^s}{a^x} < \frac{n^s}{a^{n-1}} = a \cdot \frac{n^s}{a^n} \rightarrow 0.$$

Această observaţie poate fi folosită pentru a demonstra relaţia (2). Punînd $x = \log n$ şi $e^s = a$, astfel încît $n = e^x$ şi $n^s = (e^s)^x$, cîtul din (2) devine

$$\frac{x}{a^x},$$

care este un caz particular al lui (3) pentru $s = 1$.

Exerciţii: 1). Demonstraţi că pentru $x \rightarrow \infty$, funcţia log log x tinde spre infinit mai încet decît log x. 2) Derivata funcţiei $x/\log x$ este $1/\log x - 1/(\log x)^2$. Demonstraţi că pentru valori mari ale lui x ea este „asimptotic" echivalentă cu primul termen $1/\log x$, adică raportul lor tinde spre 1 cînd $x \rightarrow \infty$.

2. Ordinul de mărime al lui log(n!)

În multe aplicaţii, ca de pildă în teoria probabilităţilor, este important să cunoaştem ordinul de mărime sau „comportarea asimptotică" a lui $n!$, pentru valori mari ale lui n. Ne vom mulţumi aici cu studiul logaritmului lui $n!$, adică al expresiei

$$P_n = \log 2 + \log 3 + \log 4 + \ldots + \log n.$$

Vom arăta că „valoarea asimptotică" a lui P_n este dată de $n \log n$, adică

$$\frac{\log (n!)}{n \log n} \rightarrow 1,$$

cînd $n \rightarrow \infty$.

Vom face demonstraţia aşa cum se obişnuieşte de obicei cînd trebuie comparată o sumă cu o integrală. În fig. 287, suma P_n este egală cu suma ariilor dreptunghiurilor ale căror baze superioare sînt trasate cu linie plină şi care la un loc nu depăşesc aria

$$\int_1^{n+1} \log x \, dx = (n+1) \log (n+1) - (n+1) + 1,$$

aflată sub curba logaritmică şi cuprinsă între 1 şi $n+1$ (cf.p. 469, exercițiul 1). Dar în același timp, suma P_n mai este egală și cu aria totală a dreptunghiurilor, ale căror baze superioare sînt trasate punctat și care la un loc depășesc aria aflată sub curbă între 1 şi n, dată de

$$\int_1^n \log x \, dx = n \log n - n + 1.$$

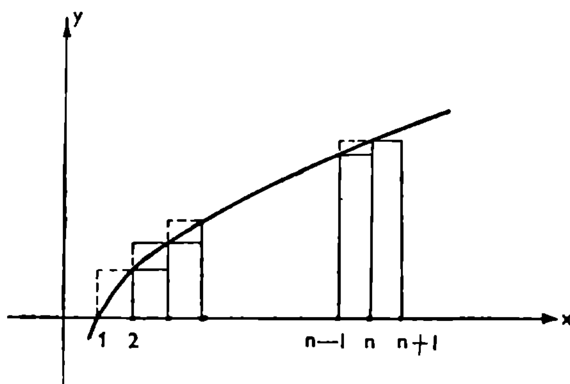


Fig. 287. Evaluarea lui log (n!)

De aici obținem

$$n \log n - n + 1 < P_n < (n+1) \log (n+1) - n.$$

și, împărțind cu $n \log n$, găsim

$$1 - \frac{1}{\log n} + \frac{1}{n \log n} < \frac{P_n}{n \log n} < (1 + 1/n) \frac{\log (n+1)}{\log n} -$$

$$- \frac{1}{\log n} = (1 + 1/n) \frac{\log n + \log (1 + 1/n)}{\log n} - \frac{1}{\log n}.$$

Evident, cele două margini tind spre 1, cînd n tinde spre infinit și afirmația noastră este demonstrată.

Exercițiu: Demonstrați că cele două margini sînt respectiv mai mare decît $1-1/n$ și mai mică decît $1 + 1/n$.

1. Serii infinite de funcţii

După cum am mai menţionat, exprimarea unei cantităţi s sub forma unei serii infinite,

(1) $$s = b_1 + b_2 + b_3 + \ldots,$$

nu este decît un mod simbolic de exprimare a faptului că s este limita şirului de „sume parţiale" finite,

$$s_1, s_2, s_3, \ldots,$$

cînd n creşte, unde

(2) $$s_n = b_1 + b_2 + \ldots + b_n.$$

Astfel, egalitatea (1) este echivalentă cu relaţia

(3) $$\lim s_n = s, \text{ cînd } n \to \infty,$$

unde s_n este definit prin (2). Cînd limita (3) există, spunem că seria (1) *converge* spre valoarea s, în timp ce dacă limita (3) nu există, spunem că seria *diverge*.

De exemplu, seria

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$$

converge spre valoarea $\pi/4$, iar seria

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

converge spre valoarea log 2; pe de altă parte, seria

$$1 - 1 + 1 - 1 + \ldots$$

diverge (deoarece sumele parţiale alternează între 1 şi 0), iar seria

$$1 + 1 + 1 + 1 + \ldots.$$

diverge, pentru că sumele parţiale tind spre infinit.

Am mai întîlnit serii ai căror termeni b_i sînt funcţii de x, de forma

$$b_i = c_i x^i,$$

cu factori constanţi c_i. Astfel de serii se numesc ***serii de puteri***; ele sînt limitele polinoamelor care reprezintă sumele parţiale

$$S_n = c_0 + c_1 x + c_2 x^2 + \dots + c_n x^n$$

(adunarea termenului constant c_0 necesită o schimbare neesenţială în notaţia (2)). O dezvoltare de forma

$$f(x) = c_0 + c_1 x + c_2 x^2 + \dots$$

a unei funcţii $f(x)$ într-o serie de puteri este astfel un mod de a aproxima funcţia $f(x)$ prin polinoame, care sînt cele mai simple funcţii. Rezumînd şi completînd rezultatele precedente, dăm următoarele dezvoltări în serii de puteri

(4) $\dfrac{1}{1+x} = 1 - x + x^2 - x^3 + \dots,$ valabilă pentru $-1 < x < +1$

(5) $\operatorname{arc\,tg} x = x - \dfrac{x^3}{3} + \dfrac{x^5}{5} - \dots,$ valabilă pentru $-1 \leqslant x \leqslant +1$

(6) $\log(1+x) = x - \dfrac{x^2}{2} - \dfrac{x^3}{3} - \dots,$ valabilă pentru $-1 < x \leqslant +1$

(7) $\dfrac{1}{2}\log\dfrac{1+x}{1-x} = x + \dfrac{x^3}{3} + \dfrac{x^5}{5} + \dots,$ valabilă pentru $-1 < x < +1$

(8) $e^x = 1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \dfrac{x^4}{4!} + \dots,$ valabilă pentru orice x.

Acestei liste îi mai adăugăm următoarele dezvoltări importante:

(9) $\sin x = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \dots,$ valabilă pentru orice x.

(10) $\cos x = 1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \dots$ valabilă pentru orice x.

Demonstraţia este o consecinţă a formulelor (vezi p. 458)

(a) $$\int_0^x \sin u\, du = 1 - \cos x,$$

(b) $$\int_0^x \cos u\, du = \sin x.$$

Începem cu inegalitatea

$$\cos x \leqslant 1.$$

Integrînd de la 0 la x, unde x este orice număr pozitiv, fixat, găsim după formula (13) de la p. 431

$$\sin x \leqslant x;$$

integrînd din nou această inegalitate, găsim

$$1 - \cos x \leqslant \frac{x^2}{2},$$

care este echivalentă cu

$$\cos x \geqslant 1 - \frac{x^2}{2}\cdot$$

Integrînd încă o dată, obținem

$$\sin x \geqslant x - \frac{x^3}{2\cdot 3} = x - \frac{x^3}{3!}\cdot$$

Procedînd mai departe în acest mod, obținem două șiruri de inegalități

$$\sin x \leqslant x \qquad\qquad \cos x \leqslant 1$$

$$\sin x \geqslant x - \frac{x^3}{3!} \qquad\qquad \cos x \geqslant 1 - \frac{x^2}{2!}$$

$$\sin x \leqslant x - \frac{x^3}{3!} + \frac{x^5}{5!} \qquad\qquad \cos x \leqslant 1 - \frac{x^2}{2!} + \frac{x^4}{4!}$$

$$\sin x \geqslant x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \qquad\qquad \cos x \geqslant 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!}$$

. . . .

Acum $x^n/n! \to 0$, cînd n tinde spre infinit. Pentru a arăta acest lucru, alegem un întreg fixat m, astfel încît $x/m < 1/2$, și scriem $c = x^m/m!$. Pentru orice întreg n, $n > m$, să punem $n = m + r$; atunci

$$0 < \frac{x^n}{n!} = c\cdot\frac{x}{m+1}\cdot\frac{x}{m+2}\cdots\frac{x}{m+r} < c\left(\frac{1}{2}\right)^r,$$

și cînd $n \to \infty$, avem și $r \to \infty$ și deci $c(1/2)^r \to 0$. Rezultă că

$$\begin{cases} \sin x = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \dfrac{x^7}{7!} + \dots \\ \\ \cos x = 1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \dfrac{x^6}{6!} + \dots . \end{cases}$$

Deoarece termenii acestor serii au semnele alternate și sînt descrescători (cel puțin pentru $|x| \leqslant 1$), rezultă că *eroarea comisă prin secționarea seriei într-un anumit loc nu depășește valoarea absolută a primului termen neglijat.*

Observații: Aceste serii pot fi folosite la calculul tabelelor.

Exemplu: Cu ce este egal sin 1°? 1° are $\pi/180$ radiani; deci

$$\sin \frac{\pi}{180} = \frac{\pi}{180} - \frac{1}{6}\left(\frac{\pi}{180}\right)^3 + \dots .$$

Eroarea comisă, dacă ne oprim aici, nu este mai mare decît $\dfrac{1}{120}\left(\dfrac{\pi}{180}\right)^5$, care este mai mic decît 0,000 000 000 02. Deci $\sin 1° = 0{,}017\,452\,406\,4$, cu zece zecimale exacte.

În sfîrșit, menționăm fără demonstrație „seria binomială".

(11) $$(1 + x)^a = 1 + ax + C_a^2 x^2 + C_a^3 x^3 + \dots,$$

unde C_a^s este „coeficientul binomial".

$$C_a^s = \frac{a(a-1)(a-2)\dots(a-s+1)}{s!}.$$

Dacă $a = n$ este un întreg pozitiv, atunci avem $C_a^n = 1$ și pentru $s > n$ toți coeficienții C_a^s din (11) sînt nuli, astfel încît obținem formula finită din teorema binomului. Una din marile descoperiri ale lui Newton, făcute la începutul carierei sale, a fost faptul că teorema binomului poate fi extinsă de la exponenții întregi pozitivi n la exponenți a arbitrari, raționali sau iraționali, pozitivi sau negativi. Dacă a nu este un întreg, atunci membrul drept din (11) dă o serie infinită, valabilă pentru $-1 < x < +1$. Pentru $|x| > 1$, seria (11) este divergentă și astfel semnul de egalitate nu are sens.

În particular, făcînd substituția $a = 1/2$ în (11), găsim dezvoltarea

(12) $$\sqrt{1 + x} = 1 + \frac{1}{2}x - \frac{1}{2^2 \cdot 2!}x^2 + \frac{1 \cdot 3}{2^3 \cdot 3!}x^3 - \frac{1 \cdot 3 \cdot 5}{2^4 \cdot 4!}x^4 + \dots .$$

Ca şi alţi matematicieni din secolul al XVIII-lea, Newton nu a dat o demonstraţie pentru formula sa. O analiză satisfăcătoare a convergenţei şi a domeniului de valabilitate ale acestor serii infinite au fost date deabia în secolul al XIX-lea.

Exerciţiu: Scrieţi seria de puteri pentru $\sqrt{1-x^2}$ şi pentru $1/\sqrt{1-x}$.

Dezvoltările (4) — (11) sînt cazuri speciale ale formulei generale a lui Brook Taylor (1685—1731), care dă dezvoltări în serii de puteri pentru o clasă largă de funcţii:

$$f(x) = c_0 + c_1x + c_2x^2 + c_3x^3 + \dots, \tag{13}$$

găsind o lege care exprimă coeficienţii c_i, cu ajutorul funcţiei f şi al derivatelor ei. Nu este posibil să dăm aici o demonstraţie riguroasă a formulei lui Taylor, enunţînd şi stabilind condiţiile valabilităţii ei. Dar consideraţiile accesibile care urmează vor lămuri, într-o oarecare măsură, corelaţia dintre faptele matematice semnificative.

Vom încerca să presupunem că este posibilă o dezvoltare de forma (13). Să mai presupunem că $f(x)$ poate fi derivată, că $f'(x)$ poate fi derivată la rîndul ei ş.a.m.d., astfel încît există un şir nesfîrşit de derivate:

$$f'(x), f''(x), \dots, f^{(n)}(x), \dots$$

În sfîrşit, admitem că o serie de puteri infinită poate fi derivată termen cu termen, ca un polinom finit. Cu aceste ipoteze putem determina coeficienţii c_n, cunoscînd comportarea lui $f(x)$ în vecinătatea lui $x = 0$. Mai întîi, făcînd $x = 0$ în (13), găsim: $c_0 = f(0)$, deoarece toţi termenii seriei care conţin pe x se anulează. Acum, derivînd egalitatea (13), obţinem:

$$f'(x) = c_1 + 2c_2x + 3c_3x^2 + \dots + nc_nx^{n-1} + \dots; \tag{13'}$$

făcînd din nou $x = 0$, dar de data aceasta în (13') şi nu în (13), găsim

$$c_1 = f'(0).$$

Derivînd egalitatea (13) obţinem:

$$f''(x) = 2c_2 + 2\cdot 3\cdot c_3x + \dots + (n-1)\cdot n\cdot c_nx^{n-2} + \dots; \tag{13''}$$

făcînd $x = 0$ în (13'') vedem că:

$$2!\, c_2 = f''(0).$$

În mod analog, derivînd egalitatea (13'') şi punînd apoi $x = 0$ găsim

$$3!\, c_3 = f'''(0),$$

şi continuînd acest procedeu, obţinem formula generală

$$c_n = \frac{1}{n!} f^{(n)}(0),$$

unde $f^n(0)$ este valoarea derivatei de ordinul n a funcţiei $f(x)$, pentru $x = 0$. Rezultatul este *seria lui Taylor*:

(14) $$f(x) = f(0) + xf'(0) + \frac{x^2}{2!} f''(0) + \frac{x^3}{3!} f'''(0) + \ldots.$$

Ca exerciţiu, cititorul poate verifica faptul că în exemplele (4) — (11) este satisfăcută legea de formare a coeficienţilor unei serii Taylor.

2. Formula lui Euler $\cos x + i \sin x = e^{ix}$

Unul dintre cele mai fascinante rezultate ale calculelor formale ale lui Euler este o legătură intimă care are loc, în domeniul numerelor complexe, între funcţiile sinus şi cosinus, pe de o parte, şi funcţia exponenţială, pe de altă parte. Ar trebui să spunem, de la început, că „demonstraţia" lui Euler şi raţionamentul nostru care urmează nu au cîtuşi de puţin un caracter riguros; ele sînt exemple tipice pentru calculele formale care se obişnuiau în secolul al XVIII-lea.

Să începem cu formula lui De Moivre, demonstrată în cap. II:

$$(\cos n\varphi + i \sin n\varphi) = (\cos \varphi + i \sin \varphi)^n.$$

Făcînd substituţia $\varphi = x/n$, obţinem formula

$$\cos x + i \sin x = \left(\cos \frac{x}{n} + i \sin \frac{x}{n}\right)^n.$$

Dacă x este dat, atunci $\cos (x/n)$ va diferi puţin de $\cos 0 = 1$ pentru valori mari ale lui n; mai mult, deoarece:

$$\frac{\sin \frac{x}{n}}{\frac{x}{n}} \to 1 \text{ cînd } \frac{x}{n} \to 0$$

(cf. p. 324), vedem că $\sin x/n$ este asimptotic egal cu x/n. De aceea considerăm plauzibilă trecerea la limită:

(14') $$\cos x + i \sin x = \lim \left(1 + \frac{ix}{n}\right)^n \text{ cînd } n \to \infty.$$

Comparînd membrul drept la acestei egalități cu formula (p. 468)

$$e^z = \lim\left(1 + \frac{z}{n}\right)^n \text{ cînd } n \to \infty$$

avem

$$\cos x + i \sin x = e^{ix}, \tag{15}$$

care este chiar rezultatul obținut de Euler.

Putem ajunge la același rezultat printr-un alt calcul formal, din dezvoltarea lui e^z:

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots,$$

în care facem substituția $z = ix$, unde x este un număr real. Dacă ne reamintim că puterile succesive ale lui i sînt i, -1, $-i$, $+1$ ş.a.m.d., în mod periodic, separînd părțile reală și imaginară, găsim:

$$e^{ix} = \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots\right).$$

Comparînd membrul drept cu seriile care dau dezvoltările funcțiilor $\sin x$ și $\cos x$, obținem din nou formula lui Euler.

Acest raționament nu este cîtuși de puțin o adevărată demonstrație a egalității (15). Obiecția pe care o putem face celui de-al doilea raționament este că dezvoltarea în serie a lui e^z a fost făcută în ipoteza că z este un număr real; de aceea, substituția $z = ix$ are nevoie de justificare. De asemenea, valabilitatea primului raționament este anulată de faptul că formula

$$e^z = \lim\left(1 + \frac{z}{n}\right)^n \text{ cînd } n \to \infty$$

a fost dedusă numai pentru valori reale ale lui z.

Pentru a scoate formula lui Euler din sfera formalismului pur și a o aduce în domeniul adevărului matematic riguros, a fost necesară dezvoltarea teoriei funcțiilor de o variabilă complexă, una dintre marile realizări matematice ale secolului al XIX-lea. Multe alte probleme au stimulat această dezvoltare cu vaste consecințe. Am văzut, de exemplu, că dezvoltările funcțiilor în serii de puteri converg pe diferite intervale. De ce unele dezvoltări converg întotdeauna, adică pentru orice x, în timp ce altele devin fără sens pentru $|x| > 1$?

Să considerăm, de pildă, seria geometrică (4) de la p. 492, care converge pentru $|x| < 1$. Membrul stîng al acestei egalități are sens cînd $x = 1$, luînd

valoarea $\frac{1}{1+1} = \frac{1}{2}$, în timp ce seria din membrul drept se comportă foarte straniu, devenind

$$1 - 1 + 1 - 1 + \ldots.$$

Această serie nu converge, deoarece sumele ei parţiale oscilează între 1 şi 0. Acest lucru indică faptul că funcţiile pot produce serii divergente, chiar dacă funcțiile nu au iregularități. Desigur, funcţia $\frac{1}{1+x}$ devine infinită cînd $x \to -1$. Deoarece se poate arăta, cu uşurinţă, că convergenţa unei serii de puteri pentru $x = a > 0$ implică întotdeauna convergenţa pentru $-a < x < a$, am putea găsi o „explicaţie" a comportării ciudate a dezvoltării în faptul că funcţia $\frac{1}{1+x}$ este discontinuă pentru $x = -1$. Totuşi funcţia $\frac{1}{1+x^2}$ poate fi dezvoltată în serie:

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \ldots$$

substituind pe x cu x^2 în (4). Această serie va converge şi ea pentru $|x| < 1$ în timp ce pentru $x = 1$ ea duce din nou la seria divergentă $1 - 1 + 1 - 1 + \ldots$, iar pentru $|x| > 1$ ea diverge brusc, cu toate că funcţia este regulată peste tot.

S-a arătat că o explicaţie completă a acestor fenomene este posibilă numai dacă funcţiile sînt studiate pentru valori *complexe* ale variabilei x, la fel ca şi pentru valorile reale. De exemplu, seria după care se dezvoltă funcţia $\frac{1}{1+x^2}$ trebuie să fie divergentă pentru $x = i$, deoarece numitorul fracţiei se anulează. Rezultă că seria trebuie să fie divergentă pentru toţi x, astfel încît $|x| > |i| = 1$, deoarece se poate arăta că convergenţa ei pentru orice astfel de număr x ar implica convergenţa ei pentru $x = i$. Astfel, problema convergenţei seriilor, complet neglijată în prima perioadă a dezvoltării analizei, a devenit unul dintre principalii factori care au contribuit la crearea teoriei funcţiilor de o variabilă complexă.

3. Seria armonică şi funcţia dzeta. Produsul lui Euler pentru sinus.

Seriile ai căror termeni sînt combinaţii simple de întregi sînt deosebit de interesante. Ca exemplu, să considerăm „seria armonică":

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n} + \ldots, \qquad (16)$$

care se deosebeşte de aceea care dă pe log 2, doar prin semnele termenilor cu numitor par.

A studia dacă această serie converge înseamnă a studia dacă şirul

$$s_1, s_2, s_3, \ldots,$$

unde

$$s_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}, \tag{17}$$

tinde spre o limită finită. Cu toate că termenii seriei (16) tind spre 0 pe măsură ce mergem mai departe în sumă, este uşor de văzut că seria nu converge. Într-adevăr, luînd destul de mulţi termeni, putem depăşi orice număr pozitiv, astfel încît s_n creşte nelimitat şi deci seria (16) „diverge spre infinit". Pentru a vedea acest lucru, observăm că:

$$s_2 = 1 + 1/2,$$

$$s_4 = s_2 + \left(\frac{1}{3} + \frac{1}{4}\right) > s_2 + \left(\frac{1}{4} + \frac{1}{4}\right) = 1 + \frac{2}{2},$$

$$s_8 = s_4 + \left(\frac{1}{5} + \ldots + \frac{1}{8}\right) > s_4 + \left(\frac{1}{8} + \ldots + \frac{1}{8}\right) =$$

$$= s_4 + \frac{1}{2} > 1 + \frac{3}{2}$$

şi, în general,

$$s_{2^m} > 1 + \frac{m}{2}. \tag{18}$$

Astfel, de exemplu, sumele parţiale s_{2^m} depăşesc pe 100, îndată ce $m \geqslant 200$.

Cu toate că seria armonică nu converge, seria

$$1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \ldots + \frac{1}{n^s} + \ldots \tag{19}$$

converge pentru orice valoare a lui s mai mare decît 1 şi defineşte pentru orice $s > 1$, aşa-numita funcţie dzeta.

$$\zeta(s) = \lim \left(1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \ldots + \frac{1}{n^s}\right) \text{ cînd } n \to \infty \tag{20}$$

ca funcţie de variabila s. Există o relaţie importantă între funcţia dzeta şi numerele prime, pe care o putem deduce folosind cunoştinţele pe care le avem referitoare la seria geometrică. Fie $p = 2, 3, 5, 7, \ldots$ un număr prim oarecare; atunci pentru $s \geqslant 1$ avem:

$$0 < \frac{1}{p^s} < 1,$$

astfel încît

$$\frac{1}{1 - 1/p^s} = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \ldots.$$

Să înmulţim aceste expresii pentru toate numerele prime $p = 2, 3, 5, 7, \ldots$, fără a ne îngriji de valabilitatea acestei operaţii. În stînga obţinem „produsul" infinit:

$$\left(\frac{1}{1 - 1/2^s}\right) \cdot \left(\frac{1}{1 - 1/3^s}\right) \cdot \left(\frac{1}{1 - 1/5^s}\right) \ldots = \text{limita pentru } n \to \infty \text{ a expresiei}$$

$$\left[\frac{1}{1 - 1/p_1^s} \cdots \frac{1}{1 - 1/p_n^s}\right],$$

în timp ce în dreapta obţinem seria:

$$1 + \frac{1}{2^s} + \frac{1}{3^s} + \ldots = \zeta(s),$$

datorită faptului că orice întreg mai mare decît 1 poate fi exprimat în mod unic, sub forma unui produs de puteri de numere prime distincte. Astfel, am reprezentat funcţia dzeta sub forma unui produs:

$$\zeta(s) = \left(\frac{1}{1 - 1/2^s}\right)\left(\frac{1}{1 - 1/3^s}\right)\left(\frac{1}{1 - 1/5^s}\right) \cdots. \qquad (21)$$

Dacă ar exista doar un număr finit de numere prime distincte, de pildă, $p_1, p_2, p_3, \ldots, p_r$, atunci produsul din membrul drept al formulei (21) ar fi un produs finit obişnuit şi ar avea deci o valoare finită şi pentru $s = 1$. Dar, după cum am văzut, seria dzeta pentru $s = 1$:

$$\zeta(1) = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$$

diverge spre infinit. Acest raţionament, care poate fi transformat cu uşurinţă într-o demonstraţie riguroasă, arată că există o infinitate de numere prime. Desigur, el este mult mai complicat şi este obţinut pe o cale mai ocolită

decît demonstrația dată de Euclid (cf. p. 39). Însă el are fascinația escaladării dificile a unui vîrf muntos, pe care am putea ajunge din partea cealaltă pe un drum confortabil.

Produse infinite ca (21) sînt uneori tot atît de folositoare pentru reprezentarea funcțiilor ca și seriile infinite. Un alt produs infinit, a cărui descoperire reprezintă încă unul dintre succesele lui Euler, se referă la funcția trigonometrică sin x. Pentru a înțelege această formulă, începem cu o observație asupra polinoamelor. Dacă $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ este un polinom de gradul n și are n zerouri distincte, $x_1, \ldots, x_n$, atunci se știe din algebră că $f(x)$ poate fi descompus în factori liniari:

$$f(x) = a_n(x - x_1) \ldots (x - x_n)$$

(cf. p. 117). Dînd factor pe $x_1 x_2 \ldots x_n$, putem scrie:

$$f(x) = C\left(1 - \frac{x}{x_1}\right)\left(1 - \frac{x}{x_2}\right) \ldots \left(1 - \frac{x}{x_n}\right),$$

unde C este o constantă egală cu a_0, după cum se poate stabili ușor, punînd $x = 0$. Acum, dacă în locul polinoamelor considerăm funcții mai complicate $f(x)$, se pune problema dacă mai este posibilă o descompunere sub forma unui produs, cu ajutorul zerourilor funcției $f(x)$. (În general, acest lucru nu poate fi adevărat, după cum se arată prin exemplul funcției exponențiale, care nu are nici un zero, deoarece $e^x \neq 0$, pentru orice valoare a lui x.) Euler a descoperit că pentru funcția sinus, o astfel de descompunere este posibilă. Pentru a scrie formula în modul cel mai simplu, considerăm nu funcția sin x, ci funcția $\sin \pi x$. Această funcție are zerourile $x = 0, \pm 1, \pm 2, \pm 3, \ldots$ deoarece $\sin \pi n = 0$ pentru orice întreg n, și numai pentru astfel de numere. Formula lui Euler afirmă că:

$$\sin \pi x = \pi x \left(1 - \frac{x^2}{1^2}\right)\left(1 - \frac{x^2}{2^2}\right)\left(1 - \frac{x^2}{3^2}\right)\left(1 - \frac{x^2}{4^2}\right) \ldots . \tag{22}$$

Acest produs infinit converge pentru toate valorile lui x și este una dintre cele mai frumoase formule din matematică. Pentru $x = 1/2$ el dă:

$$\sin \frac{\pi}{2} = 1 = \frac{\pi}{2}\left(1 - \frac{1}{2^2 \cdot 1^2}\right)\left(1 - \frac{1}{2^2 \cdot 2^2}\right)\left(1 - \frac{1}{2^2 \cdot 3^2}\right) \times$$

$$\times \left(1 - \frac{1}{2^2 \cdot 4^2}\right) \ldots .$$

Dacă scriem:

$$1 - \frac{1}{2^2 \cdot n^2} = \frac{(2n - 1)(2n + 1)}{2n \cdot 2n},$$

obținem produsul lui Wallis:

$$\frac{\pi}{2} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{6}{5} \cdot \frac{6}{7} \cdot \frac{8}{7} \cdot \frac{8}{9} \quad \ldots,$$

menționat la p. 317.

Pentru demonstrarea acestor formule trimitem cititorul la manualele de analiză (cf., de asemenea, p. 530).

** § 4. TEOREMA NUMERELOR PRIME OBȚINUTĂ PRIN METODE STATISTICE

În aplicarea metodelor matematice la studiul fenomenelor naturale ne mulțumim, de obicei, cu raționamente în cursul cărora lanțul de deducții logice riguroase este întrerupt de ipoteze mai mult sau mai puțin plauzibile. Chiar și în matematica pură întîlnim raționamente care, cu toate că nu dau o demonstrație riguroasă, sugerează totuși soluția corectă și indică direcția în care trebuie căutată demonstrația riguroasă. Soluția lui Bernoulli a problemei brahistocrone (cf. p. 400) are această trăsătură, ca și cele mai multe dintre primele lucrări făcute în analiză.

Printr-un procedeu caracteristic matematicii aplicate, și în special mecanicii statistice, vom prezenta aici un raționament care face cel puțin plauzibil adevărul celebrei legi al lui Gauss, referitoare la distribuția numerelor prime. (Un procedeu asemănător a fost sugerat unuia dintre autori, de fizicianul experimentator Gustav Hertz.) Această teoremă, discutată empiric în suplimentul cap. I, afirmă că numărul $A(n)$ de numere prime care nu depășesc pe n este asimptotic echivalent cu cantitatea $n/\log n$:

$$A(n) \sim \frac{n}{\log n}.$$

Prin aceasta înțelegem că raportul dintre $A(n)$ și $n/\log n$ tinde spre limita 1, cînd n tinde spre infinit.

Să admitem mai întîi ipoteza că *există* o lege matematică care descrie distribuția numerelor prime, în sensul următor: pentru valori mari ale lui n, funcția $A(n)$ este aproximativ egală cu integrala $\int_2^n W(x)\,\mathrm{d}x$, unde $W(x)$ este o funcție care măsoară „densitatea" numerelor prime. (Alegem pe 2 ca limită inferioară a integralei, deoarece pentru $x < 2$ avem, $A(x) = 0$.) Mai precis, fie x un număr mare și Δx un alt număr mare, astfel încît ordinul de mărime al lui x să fie mai mare decît acela al lui Δx. (De exemplu, am putea conveni

să punem $\Delta x = \sqrt{x}$.) Apoi presupunem că distribuția numerelor prime este atît de regulată, încît numărul numerelor prime din intervalul cuprins între x și $x + \Delta x$ este aproximativ egal cu $W(x) \cdot \Delta x$, și mai mult, că $W(x)$, ca funcție de x, variază atît de încet, încît integrala $\int_2^n W(x)dx$ poate fi înlocuită printr-o aproximare dreptunghiulară, fără a-i schimba valoarea asimptotică. Cu aceste observații preliminare, putem începe raționamentul.

Am demonstrat (p. 490) că pentru valori mari ale întregului n, $\log n!$ este asimptotic egal cu $n \cdot \log n$:

$$\log n! \sim n \cdot \log n.$$

Acum continuăm raționamentul, dînd o altă formulă pentru $\log n!$, în care intervin numerele prime și comparăm cele două expresii. Să numărăm de cîte ori este conținut un număr prim arbitrar p, mai mic decît n, ca factor în întregul $n! = 1 \cdot 2 \cdot 3 \ldots n$. Vom nota cu $[a]_p$ cel mai mare întreg k, pentru care p^k divide pe a. Deoarece descompunerea în factori primi a oricărui întreg este unică, rezultă că $[ab]_p = [a]_p + [b]_p$, oricare ar fi întregii a și b. Deci

$$[n!]_p = [1]_p + [2]_p + [3]_p + \ldots + [n]_p.$$

Termenii din șirul 1, 2, 3, ..., n, care sînt divizibili cu p^k, sînt p^k, $2p^k$, $3p^k$, ...; numărul lor N_k pentru n mare, este aproximativ egal cu n/p^k. Numărul M_k al acestor termeni, care sînt divizibili cu p^k, dar nu cu o putere mai mare a lui p, este egal cu $N_k - N_{k+1}$. Deci

$$\begin{aligned}[n!]_p &= M_1 + 2M_2 + 3M_3 + \ldots \\ &= (N_1 - N_2) + 2(N_2 - N_3) + 3(N_3 - N_4) + \ldots \\ &= N_1 + N_2 + N_3 + \ldots \\ &= \frac{n}{p} + \frac{n}{p^2} + \frac{n}{p^3} + \ldots = \frac{n}{p-1}.\end{aligned}$$

(Aceste egalități sînt, desigur, doar aproximative.)

Rezultă că pentru n mare, numărul $n!$ este dat în mod aproximativ de produsul tuturor expresiilor $p^{\frac{n}{p-1}}$, pentru toate numerele prime $p < n$. Obținem deci formula:

$$\log n! \sim \sum_{p<n} \frac{n}{p-1} \log p.$$

Comparînd-o cu relația asimptotică precedentă pentru $\log n!$ și punînd pe x în locul lui n, găsim

(1) $$\log x \sim \sum_{p < x} \frac{\log p}{p - 1}.$$

Pasul următor, care este și cel decisiv, este obținerea unei expresii asimptotice în funcție de $W(x)$ pentru membrul drept din (1). Cînd x este foarte mare, putem împărți intervalul de la 2 la $x = n$, într-un număr mare r de intervale parțiale mari, alegînd punctele $2 = \xi_1, \xi_2, \ldots, \xi_r, \xi_{r+1} = x$, cu creșterile corespunzătoare $\Delta\xi_j = \xi_{j+1} - \xi_j$. În fiecare interval parțial pot exista numere prime, iar toate numerele prime din cel de-al j-lea interval parțial vor avea aproximativ valoarea ξ_j. În virtutea ipotezei referitoare la $W(x)$, există aproximativ $W(\xi_j)\cdot\Delta\xi_j$ numere prime în cel de-al j-lea interval parțial; deci suma din membrul drept al lui (1) este aproximativ egală cu:

$$\sum_{j=1}^{r+1} W(\xi_j) \frac{\log \xi_j}{\xi_j - 1} \cdot \Delta\xi_j.$$

Înlocuind această sumă finită cu integrala pe care o aproximează, avem drept consecință plauzibilă a relației (1) următoarea relație:

(2) $$\log x \sim \int_2^x W(\xi) \frac{\log \xi}{\xi - 1}\, d\xi.$$

De aici vom determina funcția necunoscută $W(x)$. Dacă înlocuim semnul $\sim$ cu egalitatea obișnuită și derivăm ambii membri în raport cu x, atunci din teorema fundamentală a analizei găsim:

$$\frac{1}{x} = W(x) \frac{\log x}{x - 1},$$

$$W(x) = \frac{x - 1}{x \log x}.$$

Am presupus la începutul discuției că $A(x)$ este aproximativ egal cu $\int_2^x W(x)dx$; deci $A(x)$ este dat aproximativ de integrala:

(4) $$\int_2^x \frac{x - 1}{x \log x}\, dx.$$

Pentru a evalua această integrală, observăm că funcția $f(x) = x/\log x$ are derivata

$$f'(x) = \frac{1}{\log x} - \frac{1}{(\log x)^2}.$$

Pentru valori mari ale lui x, cele două expresii

$$\frac{1}{\log x} - \frac{1}{(\log x)^2}, \quad \frac{1}{\log x} - \frac{1}{x \log x}$$

sînt aproximaziv egale, deoarece pentru x mare al doilea termen din fiecare diferență va fi mult mai mic decît primul. Deci integrala (4) va fi asimptotic egală cu integrala:

$$\int_2^x f'(x)\,\mathrm{d}x = f(x) - f(2) = \frac{x}{\log x} - \frac{2}{\log 2},$$

deoarece integranzii vor fi aproape egali pe cea mai mare parte a domeniului de integrare. Termenul $2/\log 2$ poate fi neglijat pentru valori mari ale lui x, deoarece este o constantă, și astfel obținem rezultatul final:

$$A(x) \sim \frac{x}{\log x},$$

care este teorema numerelor prime.

Nu putem pretinde că raționamentul precedent are o valoare mai mult decît sugestivă. Dar la o analiză mai atentă apare următorul fapt: nu este greu de dat o justificare completă pentru toți pașii pe care i-am făcut cu atîta îndrăzneală; în particular, pentru relația (1), pentru echivalență asimptotică dintre această sumă și integrala din (2) și pentru pasul care duce de la (2) la (3). Este mult mai greu de demonstrat *existența* unei funcții de densitate netedă $W(x)$, pe care am presupus-o la început. O dată ce am acceptat acest lucru, *evaluarea* funcției este relativ simplă; prin urmare, în problema referitoare la numerele prime, dificultatea principală o reprezintă demonstrația existenței funcției de densitate $W(x)$.

OBSERVAȚII, PROBLEME ȘI EXERCIȚII SUPLIMENTARE

Multe dintre problemele care urmează sînt destinate cititorului mai avansat. Scopul lor nu este să dezvolte rutina de calcul, ci mai degrabă să stimuleze deprinderile inventive.

Aritmetică și algebră

1) Cum se arată că 3 nu divide nici o putere a lui 10, așa cum s-a afirmat la p. 63? (cf. p. 51).

2) Demonstrați că principiul celui mai mic întreg este o consecință a teoremei inducției matematice (cf. p. 35).

3) Aplicînd teorema binomului la dezvoltarea expresiei $(1 + 1)^n$, arătați că $C_n^0 + C_n^1 + C_n^2 + \ldots + C^n = 2^n$.

*4) Luați un întreg oarecare $z = abc\ldots$, formați suma cifrelor sale $a + b + c + \ldots$, scădeți-o din z, ștergeți o cifră din rezultat și notați cu w suma cifrelor rămase. Cunoscînd doar pe w, se poate găsi o regulă pentru determinarea valorii cifrei șterse? (Va exista un caz ambiguu, cînd $w = 0$.) Ca multe alte fapte simple, referitoare la congruențe, acesta poate fi folosit ca bază pentru un joc de societate.

5) O progresie aritmetică de ordinul întîi este un șir de numere a, $a + d$, $a + 2d$, $a + 3d$, $\ldots$, astfel încît diferența dintre doi termeni succesivi ai șirului este constantă. O progresie aritmetică de ordinul doi este un șir de numere a_1, a_2, a_3 $\ldots$, astfel încît diferențele $a_{i+1} - a_i$ formează o progresie aritmetică de ordinul întîi. În mod asemănător, o progresie aritmetică de ordinul k este un șir, astfel încît diferențele formează o progresie aritmetică de ordinul $k - 1$. Demonstrați că pătratele întregilor formează o progresie aritmetică de ordinul doi și demonstrați prin inducție că puterile k ale întregilor formează o progresie aritmetică de ordinul k. Demonstrați că orice șir al cărui termen a_n este dat de expresia $c_0 + c_1n + c_2n^2 + \ldots + c_kn^k$, unde

coeficienții c sînt constante, este o progresie aritmetică de ordinul k.** Demonstrați reciproca acestei afirmații pentru $k = 2$; $k = 3$; pentru un k oarecare.

6) Demonstrați că suma primilor n termeni ai unei progresii aritmetice de ordinul k este o progresie aritmetică de ordinul $k + 1$.

7) Cîți divizori are numărul 10 296? (cf. p. 41).

8) Folosind formula algebrică $(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$, demonstrați prin inducție că orice întreg $r = a_1 a_2 \ldots a_n$, unde toți factorii a sînt sume de două pătrate, este și el o sumă de două pătrate. Verificați această afirmație cu $2 = 1^2 + 1^2$, $5 = 1^2 + 2^2$, $8 = 2^2 + 2^2$ etc., pentru $r = 160$, $r = 1\,600$, $r = 1\,300$, $r = 625$. Dacă este posibil, dați mai multe reprezentări diferite ale acestor numere ca sumă de pătrate.

9) Aplicați rezultatul din exercițiul 8 pentru a construi noi triplete de numere pitagoreice, cu ajutorul unor triplete date.

10) Stabiliți reguli de divizibilitate, asemănătoare cu cele de la p. 51, pentru sistemele de numerație cu bazele 7, 11, 12.

11) Arătați că pentru două numere raționale pozitive $r = a/b$ și $s = c/d$, inegalitatea $r > s$ este echivalentă cu $ad - bc > 0$.

12) Arătați că pentru r și s pozitive, cu $r < s$, avem întotdeauna

$$r < \frac{r+s}{2} < s \quad \text{și} \quad \frac{2}{[(1/r) + (1/s)]^2} < 2rs < (r+s)^2.$$

13) Dacă z este un număr complex, arătați prin inducție că $z^n + 1/z^n$ poate fi exprimat ca polinom de gradul n, în raport cu cantitatea $w = z + 1/z$ (cf. p. 116).

*14) Folosind prescurtarea $\cos\varphi + i\sin\varphi = E(\varphi)$, avem: $[E(\varphi)]^m = E(m\varphi)$. Folosind acest fapt și formulele de la p. 29 referitoare la progresia geometrică, care rămîn în vigoare pentru cantități complexe, demonstrați că:

$$\sin\varphi + \sin 2\varphi + \sin 3\varphi + \ldots + \sin n\varphi = \frac{\cos\frac{\varphi}{2} - \cos\left(n + \frac{1}{2}\right)\varphi}{2\sin\frac{\varphi}{2}},$$

$$\frac{1}{2} + \cos\varphi + \cos 2\varphi + \cos 3\varphi + \ldots + \cos n\varphi = \frac{\sin\left(n + \frac{1}{2}\right)\varphi}{2\sin\frac{1}{2}\varphi}.$$

15) Ce se obține din formula din exercițiul 3 de la p. 34 dacă facem substituția $q = E(\varphi)$?

Un studiu atent al exercițiilor următoare, însoțit de desene și de exemple numerice, va contribui la stăpînirea elementelor geometriei analitice. Se presupun cunoscute definițiile și cele mai simple fapte din trigonometrie.

Adesea va fi folositor să considerăm o dreaptă sau un segment ca fiind orientat de la unul dintre punctele sale înspre un alt punct. Prin dreaptă *orientată* PQ (sau segment *orientat* în PQ), vom înțelege dreapta (sau segmentul) care are direcția de la P la Q. În lipsa unei mențiuni speciale, o dreaptă orientată l va fi presupusă avînd o orientare fixată, dar arbitrară, cu excepția faptului că axa orientată Ox va fi orientată în sensul coordonatelor pozitive și în mod asemănător pentru axa orientată Oy. Dreptele orientate (sau segmentele orientate) se vor numi atunci paralele, dacă și numai dacă ele au aceeași orientare. Orientarea unui segment orientat de pe o dreaptă orientată poate fi indicată, atașînd semnul plus sau minus distanței dintre extremitățile segmentului, după cum segmentul are aceeași orientare ca și dreapta, sau orientarea opusă. Este de dorit să extindem terminologia „segment PQ" la cazul în care P și Q coincid; un astfel de „segment" trebuie să aibă, evident, lungimea zero, dar nu va fi orientat.

16) Demonstrați că: Dacă $P_1(x_1, y_1)$ și $P_2(x_2, y_2)$ sînt două puncte oarecare, atunci coordonatele mijlocului $P_0(x_0, y_0)$ al segmentului P_1P_2 sînt $x_0 = (x_1 + x_2)/2$, $y_0 = (y_1 + y_2)/2$. Stabiliți propoziția mai generală: dacă P_1 și P_2 sînt distincte, atunci punctul P_0 de pe dreapta orientată P_1P_2, pentru care raportul $P_1P_0 : P_1P_2$ al lungimilor orientate are valoarea k, are coordonatele:

$$x_0 = (1 - k)x_1 + kx_2,$$
$$y_0 = (1 - k)y_1 + ky_2.$$

(Indicație: Dreptele paralele taie două transversale în segmente proporționale.)

Astfel punctele de pe dreapta P_1P_2 au coordonatele de forma $x = \lambda_1 x_1 + \lambda_2 x_2$, $y = \lambda_1 y_1 + \lambda_2 y_2$; cu $\lambda_1 + \lambda_2 = 1$. Valorile $\lambda_1 = 1$ și $\lambda_1 = 0$ caracterizează respectiv punctele P_1 și P_2. Valori negative ale lui λ_1 caracterizează punctele aflate dincolo de P_2, iar valorile negative ale lui λ_2 caracterizează punctele care preced pe P_1.

17) Caracterizați într-un mod asemănător poziția punctelor de pe dreaptă cu ajutorul valorilor lui k.

Tot atît de important este să folosim numerele pozitive și negative pentru a indica sensul rotațiilor, ca și al distanțelor. Prin definiție, sensul de rotație pozitiv este acela care face ca axa orientată Ox să treacă în locul axei orientate Oy, printr-o rotație de 90°. În sistemul de coordonate obișnuit, în care semiaxa Ox pozitivă este orientată spre dreapta, iar semiaxa Oy pozitivă în sus, acesta este sensul de rotație opus sensului acelor unui ceasornic. Definim acum unghiul

de la dreapta orientată l_1 la dreapta orientată l_2, ca fiind unghiul cu care trebuie să rotim dreapta l_1 pentru a face o paralelă cu l_2. Desigur, unghiul este determinat, abstracție făcînd de multiplii întregi de rotații complete, de cîte 360°. Astfel unghiul de la axa orientată Ox la axa orientată Oy este de 90° sau de —270° etc.

18) Dacă α este unghiul de la axa orientată Ox la dreapta orientată l, dacă P_1, P_2 sînt două puncte de pe l și dacă d este distanța orientată de la P_1 la P_2, arătați că

$$\cos \alpha = \frac{x_2 - x_1}{d}, \ \sin \alpha = \frac{y_2 - y_1}{d}, \ (x_2 - x_1) \sin \alpha = (y_2 - y_1) \cos \alpha.$$

Dacă dreapta l nu este perpendiculară pe axa Ox, atunci panta lui l este definită de

$$m = \operatorname{tg} \alpha = \frac{y_2 - y_1}{x_2 - x_1}.$$

Valoarea lui m nu depinde de alegerea sensului pe dreaptă, deoarece $\operatorname{tg} \alpha = \operatorname{tg} (\alpha + 180°)$ sau, ceea ce este același lucru,

$$\frac{y_1 - y_2}{x_1 - x_2} = \frac{y_2 - y_1}{x_2 - x_1}.$$

19) Demonstrați că panta unei drepte este nulă, pozitivă sau negativă, după cum o paralelă la ea, dusă prin origine, se află pe axa Ox, în primul și în al treilea cadran sau în al doilea și al patrulea cadran.

Distingem semiplanul pozitiv și cel negativ, determinat de o dreaptă orientată l, în modul următor: Fie P un punct care nu se află pe l și fie Q piciorul perpendicularei coborîte din P pe l. Atunci P este în semiplanul pozitiv sau în cel negativ determinat de l, după cum unghiul de la dreapta l la dreapta orientată QP este de 90° sau de —90°.

Vom determina acum ecuația unei drepte orientate l. Coborîm din originea O o dreaptă m perpendiculară pe l și orientăm pe m, astfel încît unghiul de la ea la l să fie de 90°. Unghiul de la axa orientată Ox la m va fi notat cu β. Atunci $\alpha = 90° + \beta$, $\sin \alpha = \cos \beta$, $\cos \alpha = -\sin \beta$. Fie $R(x_1, x_2)$ punctul în care m intersectează pe l. Vom nota cu d distanța orientată OR pe dreapta orientată m.

20) Arătați că d este pozitivă, dacă și numai dacă O se află în semiplanul negativ determinat de l.

Avem $x_1 = d \cos \beta$, $y_1 = d \sin \beta$ (cf. cu exercițiul 18).
Deci $(x - x_1) \sin \alpha = (y - y_1) \cos \alpha$ sau

$$(x - d \cos \beta) \cos \beta = -(y - d \sin \beta) \sin \beta,$$

de unde obținem ecuația

$$x \cos\beta + y \sin\beta - d = 0.$$

Aceasta este *forma normală* a ecuației dreptei l. Remarcați că această ecuație nu depinde de sensul atribuit lui l, deoarece o schimbare a sensului ar modifica semnul fiecărui termen din membrul stîng și deci ar lăsa ecuația neschimbată.

Înmulțind ecuația normală cu un factor arbitrar, obținem forma generală a ecuației dreptei:

$$ax + by + c = 0.$$

Pentru a obține din această formă generală forma normală, semnificativă din punct de vedere geometric, trebuie să o înmulțim cu un factor care să reducă primii doi coeficienți, respectiv la $\cos\beta$ și $\sin\beta$, suma pătratelor lor fiind egală cu 1. Acest lucru poate fi făcut cu ajutorul factorului $1/\sqrt{a^2+b^2}$, ceea ce ne dă forma normală

$$\frac{a}{\sqrt{a^2+b^2}}\, x + \frac{b}{\sqrt{a^2+b^2}}\, y + \frac{c}{\sqrt{a^2+b^2}} = 0,$$

astfel încît avem

$$\frac{a}{\sqrt{a^2+b^2}} = \cos\beta,\ \frac{b}{\sqrt{a^2+b^2}} = \sin\beta,\ -\frac{c}{\sqrt{a^2+b^2}} = d.$$

21) Arătați că: *a*) singurii factori care aduc forma generală la forma normală sînt $1/\sqrt{a^2+b^2}$ și $-1/\sqrt{a^2+b^2}$; *b*) alegerea unuia dintre acești factori determină orientarea dreptei; *c*) dacă unul dintre acești factori a fost ales, originea se află în semiplanul pozitiv sau în cel negativ determinat de dreapta orientată obținută, sau se află chiar pe dreaptă, după cum d este negativ, pozitiv sau nul.

22) Demonstrați direct că dreapta de pantă m, care trece printr-un punct dat $P_0(x_0, y_0)$, este dată de ecuația

$$y - y_0 = m(x - x_0) \text{ sau } y = mx + y_0 - mx_0.$$

Demonstrați că dreapta care trece prin două puncte date $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ are ecuația

$$(y_2 - y_1)(x - x_1) = (x_2 - x_1)(y - y_1).$$

Abscisa unui punct în care o dreaptă sau o curbă intersectează axa Ox se numește x — tăietura curbei. În mod asemănător se definește y — tăietura ei[4].

[4] Termenii nu sînt folosiți în literatura matematică românească, ci sînt adaptați textului. — *N.T.*

23) Împărţind ecuaţia generală din exerciţiul (20) printr-un factor convenabil, arătaţi că ecuaţia unei drepte poate fi scrisă sub forma:

$$\frac{x}{a} + \frac{y}{b} = 1,$$

numită ecuaţia prin tăieturi a dreptei, unde a şi b sînt respectiv x — şi y — tăieturi. Care sînt excepţiile?

24) Arătaţi, printr-un procedeu asemănător, că ecuaţia unei drepte care nu este paralelă cu axa Oy poate fi scrisă sub forma

$$y = mx + b.$$

(Dacă dreapta este paralelă cu axa Oy, atunci ecuaţia ei poate fi scrisă sub forma $x = a$.)

25) Fie $ax + by + c = 0$ şi $a'x + b'y + c' = 0$ ecuaţiile dreptelor neorientate l şi l', respectiv cu pantele m şi m'. Arătaţi că l şi l' sînt paralele sau perpendiculare, după cum:

a) $m = m'$ sau $mm' = -1$ b) $ab' - a'b = 0$ sau $aa' + bb' = 0$.

(Observaţi că (b) rămîne în vigoare chiar dacă dreapta nu are pantă, adică este paralelă cu axa Oy.)

26) Arătaţi că dreapta care trece printr-un punct dat $P_0(x_0, y_0)$ şi este paralelă cu o dreaptă l, de ecuaţie $ax + by + c = 0$, are ecuaţia $ax + by = ax_0 + by_0$. Arătaţi că o formulă asemănătoare, $bx - ay = bx_0 - ay_0$, este valabilă pentru ecuaţia dreptei care trece prin P_0 şi este perpendiculară pe l. (Observaţi că dacă ecuaţia l are forma normală, atunci şi noua ecuaţie va avea forma normală, în ambele cazuri.)

27) Fie $x \cos \beta + y \sin \beta - d = 0$ şi $ax + by + c = 0$, formele normală şi generală ale ecuaţiei unei drepte l. Arătaţi că distanţa orientată h de la l la orice punct Q (u, v) este dată de

$$h = u \cos \beta + v \sin \beta - d$$

sau de

$$h = \frac{au + bv + c}{\pm \sqrt{a^2 + b^2}},$$

şi că h este pozitiv sau negativ, după cum Q este în semiplanul pozitiv sau în cel negativ determinat de dreapta orientată l (sensul fiind determinat de β, sau de alegerea semnului din fața lui $\sqrt{a^2 + b^2}$). (Indicaţie: Scrieţi forma normală a ecuaţiei dreptei m care trece prin Q şi este paralelă cu l şi găsiţi distanţa de la l la m.)

28) Fie $l(x, y) = 0$ ecuaţia $ax + by + c = 0$ a unei drepte l; în mod asemănător, $l'(x, y) = 0$. Fie λ şi λ' constante cu $\lambda + \lambda' = 1$. Arătaţi că dacă l şi l' se intersectează în $P_0(x_0, y_0)$, atunci orice dreaptă care trece prin P_0 are o ecuaţie de forma

$$\lambda l(x, y) + \lambda' l'(x, y) = 0$$

şi reciproc; orice astfel de dreaptă este determinată în mod unic prin alegerea perechii de valori λ, λ'. (Indicaţie: P_0 se află pe l, dacă şi numai dacă $l(x_0, y_0) = ax_0 + by_0 + c = 0$.) Care drepte sînt reprezentate prin ecuaţia de mai sus, dacă l şi l' sînt paralele? Observaţi că condiţia $\lambda + \lambda' = 1$ nu este necesară, dar serveşte la determinarea unei ecuaţii unice pentru fiecare dreaptă care trece prin P_0.

29) Folosiţi rezultatul exerciţiului precedent pentru a găsi ecuaţia unei drepte care trece prin intersecţia P_0 a lui l şi l' şi printr-un alt punct $P_1(x_1, y_1)$, fără a găsi coordonatele lui P_0. (Indicaţie: Găsiţi pe λ şi pe λ' din condiţiile $\lambda l(x_1, y_1) + \lambda' l'(x_1, y_1) = 0$, $\lambda + \lambda' = 1$.) Verificaţi-o apoi, găsind coordonatele lui P_0 (cf. pp. 92—93) şi arătînd că P_0 se află pe dreapta a cărei ecuaţie aţi găsit-o.

30) Demonstraţi că ecuaţiile bisectoarelor unghiurilor formate de dreptele concurente l şi l' sînt:

$$\sqrt{a'^2 + b'^2}\, l(x, y) = \pm \sqrt{a^2 + b^2}\, l'(x, y).$$

(Indicaţie: a se vedea exerciţiul (27).) Ce reprezintă aceste ecuaţii, dacă l şi l' sînt paralele?

31) Găsiţi ecuaţia mediatoarei segmentului P_1P_2 prin fiecare din următoarele metode: a) Găsiţi ecuaţia dreptei P_1P_2; găsiţi coordonatele mijlocului P_0 al segmentului P_1P_2; găsiţi ecuaţia dreptei care trece prin P_0 şi este perpendiculară pe P_1P_2. b) Scrieţi ecuaţia care exprimă faptul că distanţa (cf. p. 90) dintre P_1 şi un punct oarecare $P(x, y)$ de pe mediatoare este egală cu distanţa dintre P_2 şi P; ridicaţi la pătrat ambii membri ai egalităţii şi efectuaţi reducerile.

32) Găsiţi ecuaţia cercului care trece prin trei puncte necoliniare P_1, P_2, P_3, prin fiecare din următoarele metode: a) Găsiţi ecuaţia mediatoarelor segmentelor P_1P_2 şi P_2P_3; găsiţi coordonatele centrului ca intersecţie a acestor drepte; găsiţi raza ca distanţă de la centru la P_1. b) Ecuaţia trebuie să fie de forma $x^2 + y^2 - 2ax - 2by = k$ (cf. p. 90).

Deoarece fiecare din punctele date se află pe cerc, trebuie să avem:

$$x_1^2 + y_1^2 - 2ax_1 - 2by_1 = k,$$

$$x_2^2 + y_2^2 - 2ax_2 - 2by_2 = k,$$

$$x_3^2 + y_3^2 - 2ax_3 - 2by_3 = k,$$

deoarece un punct se află pe o curbă dacă şi numai dacă coordonatele lui satisfac ecuaţia curbei. Rezolvaţi acest sistem de ecuaţii în raport cu a, b, k.

33) Pentru a găsi ecuaţia elipsei cu axa mare $2p$, axa mică $2q$ şi focarele în $F(e, 0)$ şi $F'(-e, 0)$, unde $e^2 = p^2 - q^2$, folosiţi distanţele r şi r' de la F şi F' la un punct oarecare al curbei. Din definiţia elipsei, $r + r' = 2p$. Folosind formula distanţei de la p. 90, arătaţi că:

$$r'^2 - r^2 = (x + e)^2 - (x - e)^2 = 4ex.$$

Deoarece

$$r'^2 - r^2 = (r' + r)(r' - r) = 2p(r' - r),$$

arătaţi că $r' - r = 2ex/p$. Rezolvaţi sistemul format din această relaţie şi $r' + r = 2p$, pentru a obţine formulele importante:

$$r = -\frac{e}{p}x + p, \quad r' = \frac{e}{p}x + p.$$

Deoarece $r^2 = (x - e)^2 + y^2$ (folosind din nou formula distanţei), egalaţi această expresie a lui r^2 cu expresia $\left(-\frac{e}{p}x + p\right)^2$ de mai sus:

$$(x - e)^2 + y^2 = \left(-\frac{e}{p}x + p\right)^2.$$

Dezvoltaţi, efectuaţi reducerile, substituiţi pe $p^2 - q^2$ în locul lui e^2 şi simplificaţi. Arătaţi că rezultatul poate fi exprimat sub forma:

$$\frac{x^2}{p^2} + \frac{y^2}{q^2} = 1.$$

Aplicaţi acelaşi procedeu pentru hiperbolă, definită ca loc geometric al tuturor punctelor P, pentru care valoarea absolută a diferenţei $r - r'$ este egală cu o cantitate dată $2p$. În acest caz, $e^2 = p^2 + q^2$.

34) Parabola este definită ca loc geometric al unui punct, a cărui distanţă de la o dreaptă dată (directoarea) este egală cu distanţa la un punct dat (focarul). Dacă alegem dreapta $x = -a$ ca directoare, iar punctul $F(a, 0)$ ca focar, arătaţi că ecuaţia parabolei poate fi scrisă sub forma $y^2 = 4ax$.

Construcţii geometrice

35) Demonstraţi imposibilitatea construirii cu rigla şi compasul a numerelor $\sqrt[3]{3}$, $\sqrt[3]{4}$, $\sqrt[3]{5}$. Demonstraţi că construcţia lui $\sqrt[3]{a}$ este posibilă, numai dacă a este cubul unui număr raţional (cf. p. 137 şi urm).

36) Găsiți laturile $3 \cdot 2^n$-gonului regulat și ale $5 \cdot 2^n$-gonului regulat și caracterizați șirurile corespunzătoare de extinderi ale corpurilor.

37) Demonstrați imposibilitatea trisecțiunii unui unghi de 120° sau de 30° cu ajutorul riglei și al compasului. (Indicație pentru cazul unghiului de 30°, ecuația care trebuie discutată este $4z^3 - 3z = \cos 30° = \frac{1}{2}\sqrt{3}$. Introduceți o nouă necunoscută $u = z\sqrt{3}$ și obțineți o ecuație în z, pentru care imposibilitatea construcției lui z rezultă ca în text, p. 154.)

38) Demonstrați că 9-gonul regulat nu este constructibil.

39) Demonstrați că transformarea prin inversiune a unui punct $P(x,y)$ în punctul $P'(x', y')$, față de cercul de rază r cu centrul în origine, este dată de ecuațiile

$$x' = \frac{xr}{x^2 + y^2}, \quad y' = \frac{yr}{x^2 + y^2}.$$

Găsiți pe cale algebrică ecuațiile care dau pe x, y în funcție de x', y'.

*40) Demonstrați analitic, folosind exercițiul (39), că prin inversiune mulțimea cercurilor și dreptelor este transformată în ea însăși. Verificați proprietățile a) — d) de la p. 157 în mod separat și, de asemenea, transformările care corespund fig. 61.

41) Ce se întîmplă cu cele două familii de drepte $x =$ const. și $y =$ const., paralele cu axele de coordonate, prin inversiune față de cercul de unitate cu centrul în origine? Găsiți răspunsul cu și fără ajutorul geometriei analitice (cf. p. 177).

42) Efectuați construcțiile lui Apollonius în cazuri simple, alese după voie. Verificați soluția analitic, conform metodei de la p. 142.

Geometria proiectivă și neeuclidiană

43) Găsiți toate valorile biraportului λ a patru puncte în diviziune armonică, dacă punctele sînt permutate. (Răspuns $\lambda = -1, 2, 1/2$.)

44) În care configurații formate din patru puncte coincid unele din cele șase valori ale biraportului de la p. 194? (Răspuns: numai pentru $\lambda = -1$ sau $\lambda = 1$; mai există o valoare imaginară a lui λ, pentru care $\lambda = 1/(1 - \lambda)$, biraportul „echianarmonic".)

45) Arătați că dacă biraportul $(ABCD) = 1$, atunci punctele C și D coincid.

46) Demonstrați afirmațiile referitoare la birapoartele planelor de la p. 194.

47) Demonstrați că dacă P și P' sînt inverse în raport cu un cerc și dacă diametrul AB este coliniar cu P, P', atunci punctele A, B, P, P' formează o diviziune armonică. (Indicație: folosiți expresia analitică (2) de la p. 196, luați cercul ca cerc unitate și diametrul AB ca axă.)

48) Găsiți coordonatele celui de al patrulea punct conjugat armonic față de punctele P_1, P_2, P_3. Ce se întîmplă dacă P_3 ajunge în mijlocul lui P_1P_2? (cf. p. 196).

*49) Folosiți sferele lui Dandelin pentru a dezvolta teoria secțiunilor conice. Arătați, în particular, că ele sînt toate, cu excepția cercului, locuri geometrice ale punctelor, ale căror distanțe la un punct fix F și la o dreaptă fixă l au un raport constant k. Pentru $k > 1$ avem o hiperbolă, pentru $k = 1$ o parabolă, iar pentru $k < 1$ o elipsă. Dreapta l se obține intersectînd planul conicei cu planul cercului, în lungul căruia sfera lui Dandelin atinge conul. (Deoarece cercul nu poate fi caracterizat în acest mod decît ca un caz limită, nu este chiar potrivit să alegem această proprietate ca definiție a conicelor, cu toate că este folosită uneori.)

50) Discutați propoziția: „o conică, privită atît ca mulțime de puncte, cît și ca mulțime de drepte, este autoduală" (cf. p. 228).

*51) Încercați să demonstrați teorema lui Desargues în plan, efectuînd trecerea la limită, în configurația tridimensională din fig. 73 (cf. p. 191).

*52). Cîte drepte pot fi duse astfel, încît să intersecteze patru drepte date, aflate în poziție generală în spațiu[5]? Cum pot fi caracterizate ele? (Indicație: duceți un hiperboloid prin trei din dreptele date, cf. p. 231.)

*53) Dacă cercul lui Poincaré este cercul unitate al planului complex, atunci două puncte z_1 și z_2 și valorile complexe w_1 și w_2, corespunzătoare celor două puncte de intersecție cu cercul unitate ale „dreptei" care trece prin primele două puncte definesc un biraport,

$$\frac{z_1 - w_1}{z_1 - w_2} : \frac{z_2 - w_1}{z_2 - w_2}$$

care, conform exercițiului (8) de la p. 114, este real. Logaritmul lui este, prin definiție, distanța hiperbolică dintre z_1 și z_2.

*54) Transformați cercul lui Poincaré în semiplanul superior printr-o inversiune. Dezvoltați modelul lui Poincaré și proprietățile lui pentru acest semiplan, direct și apoi cu ajutorul acestei inversiuni (cf. p. 241).

Topologie

55) Verificați formula lui Euler pentru cele cinci poliedre regulate și pentru alte poliedre. Efectuați reducerile corespunzătoare ale rețelei.

56) În demonstrația formulei lui Euler (p. 256) ni se cere să reducem orice rețea plană de triunghiuri, prin aplicarea succesivă a două operații fundamentale, la o rețea formată dintr-un singur triunghi, pentru care $V - M + F = 3 - 3 + 1 = 1$. Cum putem fi siguri că rezultatul final nu va fi o *pereche* de

[5] Spunem că un număr de drepte din spațiu se află în poziție generală, dacă două, oarecare, dintre ele nu se află în același plan. — *N.T.*

triunghiuri fără vîrfuri comune, astfel încît $V - M + F = 6 - 6 + 2 = 2$? (Indicație : putem presupune că rețeaua inițială este *conexă*, adică putem trece de la orice vîrf la oricare altul în lungul muchiilor rețelei. Arătați că această proprietate nu poate fi distrusă prin cele două operații fundamentale.)

57) Am admis doar două operații fundamentale pentru reducerea unei rețele. Nu se poate întîmpla ca, la un anumit moment, un triunghi să aibă un singur vîrf comun cu alte triunghiuri ale rețelei? (Construiți un exemplu.) Aceasta ar necesita o a treia operație : îndepărtarea a două vîrfuri, trei muchii și a unei fețe. Ar influența acest lucru demonstrația?

58) Se poate înfășura o panglică lată de cauciuc de trei ori în jurul unui băț de mătură, astfel încît ea să fie netedă (adică nerăsucită) pe bățul de mătură? (Desigur, panglica de cauciuc trebuie să se intersecteze undeva.)

59) Arătați că un disc circular din care s-a scos centrul, admite o transformare continuă în el însuși, fără puncte fixe.

*60) Transformarea care deplasează fiecare punct al unui disc cu unitatea de lungime, într-o direcție și într-un sens fixat, nu are puncte fixe. Desigur, acesta nu este o transformare a discului în *el însuși*, deoarece unele puncte vor fi scoase din disc. De ce nu rămîne în vigoare, în acest caz, raționamentul de la p. 272, bazat pe transformarea $P \rightarrow P^*$?

61) Să presupunem că avem un tub închis de cauciuc, al cărui interior este vopsit în alb, iar exteriorul este negru. Este oare posibil ca prin secționarea unei mici găuri, deformarea tubului și lipirea găurii să întoarcem tubul pe dos, astfel încît interiorul să fie negru și exteriorul alb?

*62) Arătați că în spațiu cu trei dimensiuni nu există „o problemă a celor patru culori"; pentru orice număr n putem așeza în spațiu n corpuri, astfel încît fiecare să-l atingă pe celălalt.

*63) Folosind fie suprafața unui tor (un tub închis, un inel) sau o regiune plană cu identificarea frontierei (fig. 143), construiți o hartă formată din șapte regiuni, fiecare fiind în contact cu celelalte (cf. p. 265).

64) Tetraedrul 4-dimensional din fig. 118 constă din cinci puncte, a, b, c, d, e, fiecare fiind unit cu celelalte patru. Chiar dacă curbele care le unesc pot fi încovoiate, figura nu poate fi desenată în plan, astfel încît două legături oarecare să nu se intersecteze. O altă configurație care conține zece legături și care nu poate fi trasată în plan fără intersecții constă din șase puncte, a, b, c, a', b', c', astfel încît fiecare din punctele a, b, c este legat cu fiecare din punctele a', b', c'. Verificați aceste fapte experimental și încercați să găsiți o demonstrație, folosind teorema lui Jordan. (S-a arătat că orice configurație de puncte și curbe care nu poate fi reprezentată în plan fără intersecții, trebuie să conțină una dintre aceste două configurații.)

65) Se formează o configurație, luînd cele șase muchii ale unui tetraedru 3-dimensional și adăugînd o curbă care unește mijloacele a două muchii opuse. (Două muchii ale unui tetraedru sînt opuse, dacă ele nu au nici un vîrf

comun.) Arătați că această configurație este echivalentă cu aceea descrisă în exercițiul precedent.

*66) Fie p, q, r cele trei capete ale literei E. Deplasăm litera în altă poziție, obținînd un alt E, cu vîrfurile p', q', r'. Se poate uni p cu p', q cu q' și r cu r' prin trei curbe care nu se intersectează între ele și care nu intersectează nici literele ?

Dacă ocolim un pătrat, schimbăm direcția deplasării de patru ori, de fiecare dată cu $90°$, ceea ce dă o schimbare totală de $\Delta = 360°$. Dacă ocolim un triunghi, știm din geometria elementară că $\Delta = 360°$.

67) Demonstrați că dacă C este un poligon simplu și închis atunci $\Delta = 360°$. (Indicație: secționați interiorul lui C în triunghiuri, îndepărtați apoi segmentele de la frontieră, ca la p.256.) Fie $B_1, B_2, B_3, \ldots, B_n$ frontierele succesive. Atunci $B_1 = C$, iar B_n este un triunghi. Arătați că dacă Δ_i corespunde lui B_i, atunci $\Delta_i = \Delta_{i-1}$.

68) Fie C o curbă simplă închisă, cu vector tangent continuu. Dacă Δ este variația totală a unghiului tangentei cînd parcurgem curba o dată, arătați că și aici $\Delta = 360°$.
(Indicație : fie $p_0, p_1, p_2 \ldots, p_n, p_0$ puncte care secționează curba C în segmente mici, aproape rectilinii. Fie C_i curba cu segmentele $p_0p_1, p_1p_2, \ldots p_{i-1}p_i$ și arcele inițiale $p_ip_{i+1}; \ldots p_np_0$. Atunci $C_0 = C$ iar C_n este formată din segmente rectilinii. Arătați că $\Delta_i = \Delta_{i+1}$ și folosiți rezultatul din exercițiul precedent.) Se aplică acest rezultat hipocicloidei din fig. 55?

69) Arătați că dacă în diagrama sticlei lui Klein de la p. 278, toate cele patru săgeți sînt trasate în sensul acelor unui ceasornic, se formează o suprafață echivalentă cu o sferă, în care un disc este înlocuit cu o pălărie intersectată. (Această suprafață este topologic echivalentă cu planul extins al geometriei proiective.)

70) Sticla lui Klein din fig. 142 poate fi secționată în două jumătăți simetrice printr-un plan. Arătați că fiecare din aceste părți este o bandă a lui Moebius.

*71) În banda lui Moebius din fig. 139 se identifică cele două extremități ale fiecărui segment transversal. Arătați că rezultatul este topologic echivalent cu o sticlă a lui Klein.

Toate perechile posibile de puncte de pe un segment de dreaptă, luate într-o ordine determinată (două puncte pot să și coincidă), formează un pătrat în sensul următor. Dacă notăm punctele segmentului prin distanțele lor x, y la o extremitate A, perechile ordonate de numere (x, y) pot fi privite drept coordonate carteziene ale unui punct din pătrat.

Toate perechile posibile de puncte, în care nu se ține seama de ordine (adică în care (x, y) se indentifică cu (y, x)), formează o suprafață S, topologic echivalentă cu un pătrat. Pentru a vedea acest lucru, alegeți acea reprezentare care are primul punct mai apropiat de extremitatea A a segmentului, dacă $x \neq y$. Astfel, S este mulțimea tuturor perechilor (x, y), în care x este mai mic decît

y sau $x = y$. Folosind coordonatele carteziene, aceasta dă triunghiul din plan cu vîrfurile $(0, 0), (0, 1), (1, 1)$.

*72) Ce suprafață se formează, considerînd mulțimea tuturor perechilor ordonate de puncte, din care primul aparține unei drepte, iar al doilea circumferinței unui cerc? (Răspuns: cilindru.)

73) Ce suprafață se formează, considerînd mulțimea tuturor perechilor ordonate de puncte de pe un cerc? (Răspuns: tor.)

*74) Ce suprafață se formează considerînd mulțimea tuturor perechilor *neordonate* de puncte de pe un cerc? (Răspuns: o bandă a lui Moebius.)

75) Iată regulile unui joc cu monede pe o masă circulară mare: A și B așază pe rînd monede pe masă. Monedele nu trebuie să se suprapună și fiecare poate fi așezată oriunde pe masă, fără să depășească marginea ei și fără să acopere o monedă așezată în prealabil. O dată așezată, moneda nu poate fi mișcată. Cu timpul, masa va fi acoperită cu monede, astfel încît nu mai rămîne loc pentru o altă monedă. Jucătorul care poate așeza ultima monedă pe masă cîștigă jocul. Dacă A joacă primul, demonstrați că oricum ar juca B, A poate fi sigur de cîștig, cu condiția ca el să joace corect.

76) Dacă în jocul din exercițiul (75) masa are forma din fig. 125, *b*, demonstrați că B poate cîștiga întotdeauna.

Funcții, limite și continuitate

77) Găsiți dezvoltarea în fracție continuă a cîtului $OB : AB$ de la p. 140.

78) Arătați că șirul $a_0 = \sqrt{2}$, $a_{n+1} = \sqrt{2 + a_n}$ este monoton crescător, mărginit de $B = 2$ și deci are o limită. Arătați că această limită este numărul 2 (cf. p. 142 și p. 344).

*79) Încercați să demonstrați prin metode asemănătoare cu acelea folosite la p. 336 și următoarele, că fiind dată o curbă închisă și netedă, oarecare, se poate trasa întotdeauna un pătrat, ale cărui laturi sînt tangente curbei.

Funcția $u = f(x)$ se numește *convexă*, dacă mijlocul segmentului care unește două puncte oarecare ale graficului funcției se află deasupra graficului. De exemplu, $u = e^x$ (fig. 278) este convexă, în timp ce $u = \log x$ (fig. 277) nu este.

80) Demonstrați că funcția $u = f(x)$ este convexă, dacă și numai dacă

$$\frac{f(x_1) + f(x_2)}{2} \geqslant f\left(\frac{x_1 + x_2}{2}\right),$$

egalitatea avînd loc numai pentru $x_1 = x_2$.

81) Demonstrați că pentru funcții convexe este valabilă inegalitatea mai generală

$$\lambda_1 f(x_1) + \lambda_2 f(x_2) \geqslant f(\lambda_1 x_1 + \lambda_2 x_2)$$

unde λ_1, λ_2 sînt două constante, astfel încît $\lambda_1 + \lambda_2 = 1$ şi $\lambda_1 \geqslant 0$, $\lambda_2 \geqslant 0$. Acest lucru este echivalent cu propoziţia că nici un punct al segmentului care uneşte cele două puncte ale graficului nu se află sub grafic.

82) Folosind condiţia din exerciţiul (80) demonstraţi că funcţiile $u = \sqrt{1+x^2}$ şi $u = 1/x$ (pentru $x > 0$) sînt convexe, adică

$$\frac{\sqrt{1+x_1^2}+\sqrt{1+x_2^2}}{2} \geqslant \sqrt{1+\left(\frac{x_1+x_2}{2}\right)^2},$$

$$\frac{1}{2}\left(\frac{1}{x_1}+\frac{1}{x_2}\right) \geqslant \frac{2}{x_1+x_2}$$

pentru x_1, x_2 numere întregi pozitive.

83) Acelaşi lucru pentru $u = x^2$, $u = x^n$ pentru $x > 0$, $u = \sin x$ pentru $\pi \leqslant x \leqslant 2\pi$, $u = \operatorname{tg} x$ pentru $0 \leqslant x \leqslant \pi/2$, $u = -\sqrt{1-x^2}$ pentru $|x| \leqslant 1$.

Maxime şi minime

84) Găsiţi drumul de lungime minimă între P şi Q ca în fig. 178, dacă el trebuie să întîlnească alternativ două drepte date, de n ori (cf. p. 191).

85) Găsiţi cea mai scurtă legătură între două puncte P şi Q din interiorul unui triunghi cu unghiuri ascuţite, dacă drumul trebuie să atingă laturile triunghiului într-o anumită ordine (cf. p. 352).

86) Trasaţi curbele de nivel şi verificaţi existenţa a cel puţin două puncte şa pe o suprafaţă aşezată deasupra unui domeniu triplu conex, a cărui frontieră se află la acelaşi nivel (cf. p. 353). Trebuie să excludem din nou cazul în care un plan orizontal este tangent la suprafaţă de-a lungul unei curbe închise.

87) Pornind de la două numere raţionale pozitive arbitrare a_0 şi b_0, formaţi succesiv perechile $a_{n+1} = \sqrt{a_n b_n}$, $b_{n+1} = \frac{1}{2}(a_n + b_n)$. Demonstraţi că ele determină un şir descrescător de intervale. (Punctul limită, cînd $n \to \infty$, aşa-numita medie aritmetico-geometrică a lui a şi b, a jucat un rol important în primele cercetări ale lui Gauss.)

88) Găsiţi lungimea întregului grafic din fig. 219 şi comparaţi-o cu lungimea totală a celor două diagonale.

*89) Cercetaţi condiţiile în care patru puncte A_1, A_2, A_3, A_4 duc la cazul din fig. 216 sau fig. 218.

*90) Găsiţi sisteme de cinci puncte pentu care există diferite reţele de străzi, care satisfac condiţiile unghiulare.

Doar unele vor da minime relative (cf. p. 379).

91) Demonstraţi inegalitatea lui Schwarz:

$$(a_1 b_1 + \ldots + a_n b_n)^2 \leqslant (a_1^2 + \ldots + a_n^2)(b_1^2 + \ldots + b_n^2),$$

valabilă pentru orice sistem de perechi de numere a_i, b_i; demonstrați că semnul egalității are loc numai dacă numerele a_i sînt proporționale cu b_i. (Indicație: generalizați formula algebrică din exercițiul 8.)

*92) Cu ajutorul a n numere pozitive $x_1, \ldots, x_n$ formăm expresiile s_k, definite prin

$$s_k = (x_1 x_2 \ldots x_k + \ldots)/C_n^k,$$

unde simbolul „$+\ldots$" înseamnă că trebuie adunate toate cele C_n^k produse de combinări de cîte k din aceste cantități. Atunci demonstrați că

$$\sqrt[k+1]{s_{k+1}} \leqslant \sqrt[k]{s_k},$$

unde semnul egalității are loc numai dacă toate cantitățile x_i sînt egale.

*93) Pentru $n = 3$, aceste inegalități spun că oricare ar fi numerele pozitive a, b, c, avem:

$$\sqrt[3]{abc} \leqslant \sqrt{\frac{ab + ac + bc}{3}} \leqslant \frac{a + b + c}{3}.$$

Care sînt proprietățile extremale ale cubului, implicate de aceste inegalități.

*94) Găsiți un arc de curbă de lungime minimă, care unește două puncte A, B și include cu segmentul AB o arie dată. (Răspuns: arcul trebuie să fie circular.)

*95) Fiind date două segmente AB și $A'B'$, găsiți un arc care unește pe A cu B și unul care unește pe A' cu B', astfel încît cele două arce să includă împreună cu cele două segmente o arie dată și să aibă o lungime totală minimă. (Răspuns: arcele sînt circulare și de raze egale.)

*96) Aceeași problemă pentru orice număr de segmente AB, $A'B'$ etc.

[97) Pe două drepte care se întîlnesc în O, găsiți două puncte A și B și uniți pe A cu B printr-un arc de lungime minimă, astfel încît aria inclusă de el și de dreptele date să fie dată. (Răspuns: arcul este circular și ortogonal pe cele două drepte date.)

*98) Aceeași problemă, însă acum perimetrul total al domeniului, adică arcul plus OA plus OB, trebuie să fie minim. (Răspuns: soluția este dată de un arc de cerc, care iese în afară și este tangent celor două drepte.)

*99) Aceeași problemă, pentru mai multe sectoare unghiulare.

*100) Demonstrați că suprafețele aproape plane din fig. 240 nu sînt plane, cu excepția suprafeței stabilizate în centru. Observație: a găsi sau a caracteriza analitic aceste suprafețe este o problemă încă nerezolvată. Același lucru este adevărat pentru suprafețele din fig. 251. În fig. 258 avem douăsprezece plane simetrice, care se întîlnesc în lungul diagonalelor, făcînd unghiuri de 120°.

Alte experiențe cu pelicule de săpun. Efectuați experiențele indicate de fig. 256 și fig. 257 pentru cazul a mai mult de trei bare de legătură. Studiați cazurile limită, cînd volumul aerului tinde spre zero. Experimentați cu plane neparalele sau cu alte suprafețe. Suflați în balonul cubic din fig. 258, pînă ce el umple întregul cub și apoi îl depășește. Aspirați apoi aerul, inversînd procesul.

*101) Găsiți două triunghiuri echilaterale cu perimetrul total dat și cu arie minimă. (Răspuns: triunghiurile trebuie să fie egale. Folosiți metodele calculului diferențial.)

*102) Găsiți două triunghiuri cu perimetrul total dat și cu arie maximă. (Răspuns: Un triunghi degenerează într-un punct iar celălalt trebuie să fie echilateral.)

*103) Găsiți două triunghiuri cu arie totală dată și cu perimetru minim.

104) Găsiți două triunghiuri echilaterale, cu arie totală dată și cu perimetru maxim.

Calculul diferențial și integral

105) Derivați funcțiile $\sqrt{1+x}$, $\sqrt{1+x^2}$, $\sqrt{\dfrac{x+1}{x-1}}$ folosind direct definiția derivatei, formînd și transformînd cîtul creșterilor, pînă ce limita poate fi obținută cu ușurință făcînd substituția $x_1 = x$ (cf. p. 439).

106) Demonstrați că funcția $y = e^{-1/x^2}$, cu $y = 0$ pentru $x = 0$, are toate derivatele nule în $x = 0$.

107) Arătați că funcția din exercițiul (106) nu poate fi dezvoltată în serie Taylor (cf. p. 496).

108) Găsiți punctele de inflexiune ($f''(x) = 0$) ale curbelor $y = e^{-x^2}$ și $y = xe^{-x^2}$.

109) Demonstrați că pentru un polinom $f(x)$ cu toate cele n rădăcini $x_1, \ldots, x_n$ distincte, avem:

$$\frac{f'(x)}{f(x)} = \sum_{i=1}^{n} \frac{1}{x - x_i}.$$

*110) Pornind de la definiția integralei ca limită a unei sume, demonstrați că pentru $n \to \infty$, avem

$$n\left(\frac{1}{1^2+n^2} + \frac{1}{2^2+n^2} + \ldots + \frac{1}{n^2+n^2}\right) \to \frac{\pi}{4}.$$

*111) Demonstrați în mod asemănător că

$$\frac{b}{n}\left(\sin\frac{b}{n} + \sin\frac{2b}{n} + \ldots + \sin\frac{nb}{n}\right) \to \cos b - 1.$$

112) Desenînd fig. 276 la o scară mai mare, pe hîrtie milimetrică şi numărînd micile pătrate din aria haşurată, găsiţi o valoare aproximativă a lui π.

113) Folosiţi formula (7) de la p. 460, pentru calculul numeric al lui π, cu o eroare de cel mult 1/100.

114) Demonstraţi că $e^{\pi i} = -1$ (cf. p. 498).

115) O curbă de formă dată se dilată în raportul $1 : x$. Fie $L(x)$ şi $A(x)$ respectiv lungimea şi aria curbei obţinute. Arătaţi că $L(x)/A(x) \to 0$ cînd $x \to \infty$, şi, mai general, $L(x)/A(x)^k \to 0$ cînd $x \to \infty$, dacă $k > 1/2$. Verificaţi pentru cerc, pătrat şi elipsă. (Aria are un ordin de mărime superior lungimii circumferinţei; cf. p. 492.)

116) Funcţia exponenţială apare adesea în combinaţii de tipul indicat mai jos şi notate aşa cum se arată:

$$u = \text{sh } x = \frac{1}{2}(e^x - e^{-x}), \quad v = \text{ch } x = \frac{1}{2}(e^x + e^{-x})$$

$$w = \text{th } x = \frac{e^x - e^{-x}}{e^x + e^{-x}},$$

numite respectiv *sinusul hiperbolic*, *cosinusul hiperbolic* şi *tangenta hiperbolică*. Aceste funcţii au multe proprietăţi care ne amintesc de proprietăţile funcţiilor trigonometrice; ele sînt legate de hiperbola $u^2 - v^2 = 1$, tot aşa cum funcţiile $u = \cos x$ şi $v = \sin x$ sînt legate de cercul $u^2 + v^2 = 1$. Cititorul ar trebui să demonstreze următoarele proprietăţi şi să le compare cu proprietăţile corespunzătoare referitoare la funcţiile trigonometrice

$$D \text{ ch } x = \text{sh } x, \qquad D \text{ sh } x = \text{ch } x, \qquad D \text{ tgh } x = 1/\text{ch}^2 x,$$

$$\text{sh}(x + x') = \text{sh } x \cdot \text{ch } x' + \text{ch } x \cdot \text{sh } x',$$

$$\text{ch}(x + x') = \text{ch } x \cdot \text{ch } x' + \text{sh } x \cdot \text{sh } x'.$$

Funcţiile inverse se numesc $x = \text{arg sh } u = \log(u + \sqrt{u^2 + 1})$; $x = \text{arg ch } v = \log(v + \sqrt{v^2 - 1})$ $(v \geq 1)$.

Derivatele lor sînt date de:

$$D \text{ arg sh } u = \frac{1}{\sqrt{1 + u^2}}; \; D \text{ arg ch } v = \frac{1}{\sqrt{v^2 - 1}} \; D \text{ arg th } w = \frac{1}{1 - w^2} \; (|w| > 1).$$

117) Folosind formula lui Euler, verificaţi analogia dintre funcţiile hiperbolice şi cele trigonometrice.

*118) Găsiţi formule de sumare simple pentru:

$$\text{sh } x + \text{sh } 2x + \ldots + \text{sh } nx$$

şi pentru

$$\frac{1}{2} + \text{ch } x + \text{ch } 2x + \ldots + \text{ch } nx$$

analoge cu acelea din exercițiul (14) pentru funcțiile trigonometrice.

Tehnica integrării

Teorema demonstrată la p. 457 reduce problema integrării unei funcții $f(x)$, între limitele a și b, la aceea a găsirii unei funcții primitive $G(x)$ a funcției $f(x)$, adică a unei funcții astfel, încît $G'(x) = f(x)$. Integrala este atunci egală cu diferența $G(b) - G(a)$. Pentru aceste funcții primitive, care sînt determinate de funcția $f(x)$ (cu excepția unei constante aditive), se folosește notația sugestivă

$$G(x) = \int f(x)\mathrm{d}x,$$

în care nu apar limitele de integrare. (Această notație poate fi derutantă pentru începător; a se vedea observația de la p. 456.)

Orice formulă de derivare conține soluția unei probleme de integrare prin simpla ei citire inversă ca formulă de integrare. Putem extinde acest procedeu oarecum empiric cu ajutorul a două reguli importante, care nu sînt decît echivalentul regulilor de derivare a funcției compuse și a unui produs de funcții. Sub forma lor integrală, ele se numesc regulile *integrării prin substituție* și a *integrării prin părți*.

A) Prima regulă rezultă din formula de derivare a unei funcții compuse,

$$H(u) = G(x),$$

unde

$$x = \psi(u) \text{ și } u = \varphi(x)$$

sînt presupuse funcții una de cealaltă, determinate în mod unic în intervalul considerat. Atunci avem:

$$H'(u) = G'(x)\,\psi'(u).$$

Dacă

$$G'(x) = f(x),$$

putem scrie

$$G(x) = \int f(x)\,\mathrm{d}x$$

și

$$G'(x)\psi'(u) = f(x)\psi'(u),$$

ceea ce, ținînd seama de formula de mai sus pentru $H'(u)$, este echivalent cu

$$H(u) = \int f(\psi(u))\psi'(u)\mathrm{d}u.$$

Deci, deoarece $H(u) = G(x)$,

$$\text{I)} \qquad \int f(x)\,\mathrm{d}x = \int f(\psi(u))\psi'(u)\mathrm{d}u.$$

Cu notația lui Leibniz (cf. p. 460), această regulă capătă forma sugestivă

$$\int f(x)\,\mathrm{d}x = \int f(x)\frac{\mathrm{d}x}{\mathrm{d}u}\mathrm{d}u,$$

ceea ce înseamnă că simbolul $\mathrm{d}x$ poate fi înlocuit prin simbolul $\frac{\mathrm{d}x}{\mathrm{d}u}\mathrm{d}u$, ca și cum $\mathrm{d}x$ și $\mathrm{d}u$ ar fi numere și $\frac{\mathrm{d}x}{\mathrm{d}u}$ o fracție.

Utilitatea formulei (I) va fi ilustrată prin cîteva exemple:

a) $J = \int \frac{1}{u \log u}\mathrm{d}u$. Începem cu membrul drept al formulei (I), făcînd substituția $x = \log u = \psi(u)$. Atunci obținem

$$\psi'(u) = \frac{1}{u},\ f(x) = \frac{1}{x};\ \text{deci}$$

$$J = \int \frac{\mathrm{d}x}{x} = \log x$$

sau

$$\int \frac{\mathrm{d}u}{u \log u} = \log \log u.$$

Putem verifica acest rezultat, derivînd ambii membri. Găsim $\frac{1}{u \log u} = \frac{\mathrm{d}}{\mathrm{d}u}(\log \log u)$, ceea ce se vede cu ușurință că este adevărat.

b) $J = \int \operatorname{ctg} u\,\mathrm{d}u = \int \frac{\cos u}{\sin u}\mathrm{d}u$. Punînd $x = \sin u = \psi(u)$, găsim

$$\psi'(u) = \cos u,\ f(x) = x,$$

de unde rezultă

$$J = \int \frac{dx}{x} = \log x$$

sau

$$\int \operatorname{ctg} u \, du = \log \sin u.$$

Acest rezultat poate fi verificat din nou prin derivare.

c) În general, dacă avem o integrală de forma:

$$J = \int \frac{\psi'(u)}{\psi(u)} \, du,$$

punem $x = \psi(u)$, $f(x) = x$ și găsim

$$J = \int \frac{dx}{x} = \log x = \log \psi(u).$$

d) $J = \int \sin x \cos x \, dx$. Punem $\sin x = u$, $\cos x = \dfrac{du}{dx}$.

Atunci

$$J = \int u \frac{du}{dx} \, dx = \int u \, du = \frac{u^2}{2} = \frac{1}{2} \sin^2 x.$$

e) $J = \displaystyle\int \frac{\log u}{u} \, du.$

Punem $\log u = x$, $\dfrac{1}{u} = \dfrac{dx}{du}$. Atunci

$$J = \int x \frac{dx}{du} \, du = \int x \, dx = \frac{x^2}{2} = \frac{1}{2} (\log u)^2.$$

În exemplele următoare se folosește formula (I), pornind din membrul stîng.

f) $J = \displaystyle\int \frac{dx}{\sqrt{x}}$. Punem $\sqrt{x} = u$. Atunci $x = u^2$ și $\dfrac{dx}{du} = 2u$.

De aceea:

$$J = \int \frac{1}{u} \cdot 2u \; du = 2u = 2\sqrt{x}.$$

g) Făcînd substituția $x = au$, unde a este o constantă, găsim

$$\int \frac{dx}{a^2 + x^2} = \int \frac{dx}{du} \cdot \frac{1}{a^2} \cdot \frac{1}{1 + u^2} du = \int \frac{1}{a} \cdot \frac{du}{1 + u^2} = \frac{1}{a} \operatorname{arc\,tg} \frac{x}{a}.$$

h) $J = \int \sqrt{1 - x^2}\, dx$. Să punem $x = \cos u$, $\frac{dx}{du} = -\sin u$.

În acest caz,

$$J = -\int \sin^2 u \, du = -\int \frac{1 - \cos 2u}{2} du = -\frac{u}{2} + \frac{\sin 2u}{4}.$$

Folosind formula $\sin 2u = 2 \sin u \cos u = 2 \cos u \sqrt{1 - \cos^2 u}$,

avem

$$J = -\frac{1}{2} \operatorname{arc\,cos} x + \frac{1}{2} x \sqrt{1 - x^2}.$$

Calculați următoarele primitive și verificați rezultatele prin derivare

119) $\int \frac{u\, du}{u^2 - u + 1}$.

120) $\int u e^{u^2} du$.

121) $\int \frac{du}{u(\log u)^n}$.

122) $\int \frac{8x}{3 + 4x} dx$.

123) $\int \frac{dx}{x^2 + x + 1}$.

124) $\int \frac{dx}{x^2 + 2ax + b}$.

125) $\int t^2 \sqrt{1 + t^3}\, dt$.

126) $\int \frac{t + 1}{\sqrt{1 - t^2}} dt$.

127) $\int \frac{t^4}{1 - t} dt$.

128) $\int \cos^n t \cdot \sin t \cdot dt$.

129) Demonstrați că $\int \frac{dx}{a^2 - x^2} = \frac{1}{a} \operatorname{arg\,th} \frac{x}{a}$; $\int \frac{dx}{\sqrt{a^2 + x^2}} = \operatorname{arg\,sh} \frac{x}{a}$.
(Comparați cu exemplele $(g)(h)$.)

B) Regula pentru derivarea unui produs (p. 446)

$$(p(x) \cdot q(x))' = p(x) \cdot q'(x) + p'(x) \cdot q(x),$$

poate fi scrisă ca formulă integrală, sub forma:

$$p(x) \cdot q(x) = \int p(x) \cdot q'(x) \mathrm{d}x + \int p'(x) \cdot q(x) \mathrm{d}x$$

sau

$$\text{(II)} \qquad \int p(x)q'(x)\mathrm{d}x = p(x)q(x) - \int p'(x)q(x)\mathrm{d}x.$$

Sub această formă ea se numeşte regula de *integrare prin părți*. Această regulă este folositoare în cazul în care funcția care trebuie integrată poate fi scrisă ca produs de forma $p(x)q'(x)$, unde este cunoscută funcția primitivă $q(x)$ a funcției $q'(x)$. În acest caz, formula (II) reduce problema găsirii primitivei funcției $p(x)q'(x)$ la aceea a găsirii primitivei funcției $p'(x)q(x)$, care adesea este mult mai uşor de rezolvat.

Exemple:

a) $J = \int \log x \, \mathrm{d}x$. Puneți $p(x) = \log x$, $q'(x) = 1$, astfel încît $q(x) = x$.

Atunci formula (II) ne conduce la

$$\int \log x \, \mathrm{d}x = x \log x - \int \frac{x}{x} \mathrm{d}x = x \log x - x.$$

b) $J = \int x \log x \, \mathrm{d}x$. Punem $p(x) = \log x$, $q'(x) = x$.

Atunci

$$J = \frac{x^2}{2} \log x - \int \frac{x^2}{2x} \mathrm{d}x = \frac{x^2}{2} \log x - \frac{x^2}{4} \cdot$$

c) $J = \int x \sin x \, \mathrm{d}x$. De data aceasta punem $p(x) = x$, $q(x) = -\cos x$ şi găsim

$$\int x \sin x \, \mathrm{d}x = -x \cos x + \sin x.$$

Calculați următoarele primitive, folosind integrarea prin părți.

130) $\int xe^x \, dx$.

131) $\int x^2 \cos x \, dx$.

(Indicație: Aplicați formula (II) de două ori.)

132) $\int x^a \log x \, dx \; (a \neq -1)$.

133) $\int x^2 e^x \, dx$.

(Indicație: folosiți exercițiul 130.)

Integrarea prin părți a integralei $\int \sin^m x \, dx$ duce la o expresie remarcabilă a numărului π, sub forma unui produs infinit. Pentru a o deduce, să scriem funcția $\sin^m x$ sub forma $\sin^{m-1} x \cdot \sin x$ și să integrăm prin părți între limitele 0 și $\pi/2$. Aceasta duce la formula

$$\int_0^{\pi/2} \sin^m x \, dx = (m-1) \int_0^{\pi/2} \sin^{m-2} x \cos^2 x \, dx =$$

$$= -(m-1) \int_0^{\pi/2} \sin^m x \, dx + (m-1) \int_0^{\pi/2} \sin^{m-2} x \, dx,$$

sau

$$\int_0^{\pi/2} \sin^m x \, dx = \frac{m-1}{m} \int_0^{\pi/2} \sin^{m-2} x \, dx,$$

deoarece primul termen din membrul drept al formulei (II), pq este egal cu 0 pentru valorile 0 și $\pi/2$. Prin aplicarea repetată a ultimei formule, găsim următoarea valoare pentru $I_m = \int_0^{\pi/2} \sin^m x \, dx$ (formulele diferă, după cum m este par sau impar):

$$I_{2n} = \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} \cdots \frac{1}{2} \cdot \frac{\pi}{2},$$

$$I_{2n+1} = \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} \cdots \frac{2}{3}.$$

Deoarece $0 < \sin x < 1$ pentru $0 < x < \pi/2$, avem $\sin^{2n-1} x > \sin^{2n} x > \sin^{2n+1} x$ şi, prin urmare,

$$I_{2n-1} > I_{2n} > I_{2n+1} \qquad \text{(cf. p. 431)}$$

sau

$$\frac{I_{2n-1}}{I_{2n+1}} > \frac{I_{2n}}{I_{2n+1}} > 1.$$

Substituind valorile calculate mai sus ale lui I_{2n-1}, în ultimele inegalităţi, găsim

$$\frac{2n+1}{2n} > \frac{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \ldots (2n-1)(2n-1)(2\,n+1)}{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \ldots (2n)(2n)} \cdot \frac{\pi}{2} > 1.$$

Dacă trecem acum la limită, făcînd $n \to \infty$, vedem că termenul din mijloc tinde spre 1 şi deci obţinem reprezentarea lui Wallis a numărului $\pi/2$ sub forma unui produs:

$$\frac{\pi}{2} = \frac{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot \ldots 2n \cdot 2n \ldots}{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \ldots (2n-1) \cdot (2n-1) \cdot (2n+1) \ldots} =$$

$$= \lim \frac{2^{4n}(n\,!)^4}{[(2n)\,!]^2\,(2n+1)} \quad \text{cînd } n \to \infty.$$

SUGESTII PENTRU UN STUDIU ULTERIOR

Referințe generale

AHRENS, W., *Mathematische Unterhaltungen und Spiele*, ed. a II-a, 2 vol. Leipzig, Teubner, 1910.
BALL, W. W. ROUSE, *Mathematical Recreations and Essays*, ed. a XI-a revăzută de H.S.M. Coxeter, New York, MacMillan, 1939.
BELL, E.T., *The Development of Mathematics*, New York, McGraw-Hill, 1940.
IDEM, *Men of Mathematics*, New York, Simon and Schuster, 1937.
DANTZIG, T., *Aspects of Science*, New York, MacMillan, 1937.
DRESDEN, A., *An Invitation to Mathematics*, New York, Holt, 1936.
ENRIQUES, F., *Questioni riguardanti le matematiche elementari*, ed. a III-a, 2 vol. Bologna, Zanichelli, 1924 și 1926.
KASNER, E. și NEWMAN, J., *Mathematics and the Imagination*, New York, Simon and Schuster, 1940.
KLEIN, F., *Elementary Mathematics from an Advanced Standpoint*, tradusă de E.R. Hedrick și C. A. Nohle, 2. vol., New York, McMillan, 1932 și 1939.
KRAITCHIK, M., *La Mathématique des Jeux*. Bruxelles, Stevens, 1930.
NEUGEBAUER, O., *Vorlesungen über Geschichte der antiken matematischen Wissenschaften*, Erster Band: *Vorgriechische Mathematik*, Berlin, Springer, 1934.
POINCARÉ, H., *The Foundations of Science*, Lancaster, Pa: Science Press, 1913.
RADEMACHER, H. și TOEPLITZ, O., *Von Zahlen und Figuren*, ed. a II-a, Berlin, Springer, 1933. (Există traducerea în limba română, *Despre numere și figuri*, Editura Științifică, București, 1968.)
RUSSELL, B., *Introduction to Mathematical Philosophy*, London, Allen and Unwin, 1924.
IDEM, *The Principles of Mathematics*, ed. a II-a, New York, Norton, 1938.
SMITH, D. E., *A Source Book in Mathematics*, New York, McGraw-Hill, 1929.
STEINHAUS, H., *Mathematical Snapshots*, New York, Stechert, 1938. (Există traducerea în limba română: *Caleidoscop matematic*, Editura Tehnică, București.)
WEYL, H., *The Mathematical Way of Thinking*, în „Science", XCII, 1940, p. 437 și urm.
IDEM, *Philosophie der Mathematik und Naturwissenschaft*, Handbuch der Philosophie, Band II, München: Oldenbourg, 1926, pp. 3—162.

Capitolul I

DICKSON, L. E., *Introduction to the Theory of Numbers*, Chicago, University of Chicago Press, 1931.
IDEM, *Modern Elementary Theory of Numbers*, Chicago, University of Chicago Press, 1939.

HARDY, G. H., *An Introduction to the Theory of Numbers*, în „Bulletin of the American Mathematical Society", XXXV, 1929, p. 789 și urm.
HARDY, G. H. și WRIGHT, E. M., *An Introduction of the Theory of Numbers*, Oxford, Clarendon Press, 1938.
USPENSKY, J. V. și HEASLET, M. H., *Elementary Number Theory*, New York, McGrav-Hill, 1939.

Capitolul II

BIRKHOFF, G. și MacLANE, S., *A Survey of Modern Algebra*, New York, McMillan, 1941.
BLACK, M., *The Nature of Mathematics*, New York, Harcourt, Brace, 1935.
DANTZIG, T., *Number, the Language of Science*, ed. a III-a, New York, McMillan, 1939.
HARDY, G. H., *A Course of Pure Mathematics*, ed. a VII-a, Cambridge, University Press, 1938.
KNOPP, K., *Theory and Application of Infinite Series*, tradusă de Miss R.C. Young, Londra, Blackie, 1928.
TARSKI, A., *Introduction to Logic*, New York, Oxford, University Press, 1939.
ENRIQUES, F., *The History Development of Logic*, tradusă de J. Rosenthal, New York, Holt, 1929.

Capitolul III

COOLIDGE, J. L., *A History of Geometrical Methods*, Oxford, Clarendon Press, 1940.
DICKSON, L. E., *New First Course in the Theory of Equations*, New York, Willey, 1939.
ENRIQUES, F. (editor), *Fragen der Elementargeometrie*, ed. a II-a, 2 vol., Leipzig, Teubner, 1923.
HOBSON, E. W., *„Squaring the Circle", A History of the Problem*, Cambridge, University Press, 1913.
KEMPE, A. B., *How to Draw a Straight Line*, London, MacMillan, 1877.
KLEIN, F., *Famous Problems of Geometry*, tradusă de W. W. Beman și D. E. Smith, ed. a II-a, New York, Stechert, 1930.
MASCHERONI, L., *La geometria del compasso*, Palermo, Reber, 1901.
MOHR, G., *Euclides Danicus*, Copenhaga, Holst, 1928.
MORGAN, A., de, *A Budget of Paradoxes*, 2 vol., Chicago, Open Court, 1915.
THOMAS, J. M., *Theory of Equations*, New York, McGraw-Hill, 1938.
WEISNER, L., *Introduction to the Theory of Equations*, New York, Willey, 1939.

Capitolul IV

GRAUSTEIN, W. C., *Introduction to Higher Geometry*, New York, McMillan, 1930.
HILBERT, D., *The Foundations of Geometry*, tradusă de E. J. Townsend, ed. a III-a, La Salle, III: Open Court, 1938.
O'HARA, C. W. și WARD, D. R., *An Introduction to Projective Geometry*, Oxford, Clarendon Press, 1937.
ROBINSON, G. B., de, *The Foundations of Geometry*, Toronto, University of Toronto Press, 1940.
SACCHERI, G., *Euclides ab omni naevo vindicatus*, tradusă de G. B. Halsted, Chicago, Open Court, 1920.
SANGER, R. G., *Synthetic Projective Geometry*, New York, McGraw-Hill, 1939.
VEBLEN, O. și YOUNG, J. W., *Projective Geometry*, 2 vol., Boston, Ginn, 1910 și 1918.
YOUNG, J. W., *Projective Geometry*, Chicago, Open Court, 1930.

Capitolul V

ALEXANDROFF, P., *Einfachste Grundbegriffe der Topologie*, Berlin, Springer, 1932.
HILBERT, D. şi COHN-VOSSEN, S., *Anschauliche Geometrie*, Berlin, Springer, 1932.
NEWMAN, M. H. A., *Elements of the Topology of Plane Sets of Points*, Cambridge, University Press, 1939.
SEIFERT, H. şi THRELFALL, W., *Lehrbuch der Topologie*, Leipzig, Teubner, 1934.

Capitolul VI

COURANT, R., *Differential and Integral Calculus*, tradusă de E. J. McShane, ediţie revăzută, 2 vol., New York, Nordemann, 1940.
HARDY, G. H., *A Course of Pure Mathematics*, ed. a VII-a, Cambridge, University Press, 1938.
FERRAR, W. L., *A Text-book of Convergence*, Oxford, Clarendon Press, 1938.
Pentru teoria fracţiilor continue, a se vedea de exemplu:
BARNARD, S. şi CHILD, J. M., *Advanced Algebra*, London, McMillan, 1939.

Capitolul VII

COURANT, R., *Soap Film Experiments with Minimal Surfaces*, în „American Mathematical Monthly", XLVII, 1940, pp. 167—174.
PLATEAU, J., *Sur les figures d'équilibre d'une masse liquide sans pésanteur*, în „Mémoires de l'Académie Royale de Belgique", Nouvelle Série, XXIII (1849).
IDEM, *Statique expérimentale et théoretique des Liquides*, Paris, 1873.

Capitolul VIII

BOYER, C. B., *The Concepts of the Calculus*, New York, Columbia, University Press, 1939.
COURANT, R., *Differential and Integral Calculus*, tradusă de E.J. McShane, ediţie revăzută, 2 vol., New York, Nordemann, 1940.
HARDY, G. H., *A Course of Pure Mathematics*, ed. a VII-a, Cambridge, University Press, 1938.

INDICE

D

E

F

G

H

I, Î

L

M

Ș

T

U

V

Z

CUPRINSUL

Capitolul I

NUMERELE NATURALE

Supliment la capitolul I

TEORIA NUMERELOR

Capitolul II

SISTEMUL DE NUMERE AL MATEMATICII

Supliment la capitolul II

ALGEBRA MULȚIMILOR

Capitolul III

CONSTRUCȚII GEOMETRICE. ALGEBRA CÎMPURILOR DE NUMERE

Capitolul IV

GEOMETRIA PROIECTIVĂ. AXIOMATICA. GEOMETRII NEEUCLIDIENE

Apendice

Capitolul V

TOPOLOGIA

Apendice

Capitolul VI

FUNCȚII ȘI LIMITE

Supliment la capitolul VI

ALTE EXEMPLE REFERITOARE LA LIMITE ȘI CONTINUITATE

Capitolul VII

MAXIME ȘI MINIME

Capitolul VIII

CALCULUL DIFERENȚIAL ȘI INTEGRAL

Supliment la capitolul VIII

Apendice

OBSERVAȚII, PROBLEME ȘI EXERCIȚII SUPLIMENTARE